

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 777.075 Seconds
(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-1

Sequence: 1 CGGGATCCCTTAGAATATTA.....TTAAGCGAGCGAATTCGG 100

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_ets:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	100	6	BD006932 Artificial
2	41	41.0	113	6	BD006933 Artificial
3	24	24.0	55730	2	ACT02301 Mus muscu
4	24	24.0	128986	2	ACT02301 Mus muscu
5	24	24.0	128986	2	ACT02301 Mus muscu
6	23	23.0	63125	2	ACT02301 Mus muscu
7	23	23.0	110000	2	ACT02301 Mus muscu
8	23	23.0	164840	2	ACT02301 Mus muscu
9	23	23.0	175654	2	ACT02301 Mus muscu
10	23	23.0	210911	2	ACT02301 Mus muscu
11	23	23.0	212722	2	ACT02301 Mus muscu
12	23	23.0	230777	2	ACT02301 Mus muscu
13	23	23.0	247062	2	ACT02301 Mus muscu
14	22	22.0	1384	6	CO493603 Sequence
15	22	22.0	52615	2	AC090029 Arabidops
16	22	22.0	59935	2	AC090029 Arabidops
17	22	22.0	63933	2	AC090029 Arabidops
18	22	22.0	64118	2	AC090029 Arabidops
19	22	22.0	65934	2	AC126380 Homo sapi

20	22	22.0	67276	2	AC026622 Homo sapi
21	22	22.0	69183	2	AC123623 Mus muscu
22	22	22.0	70089	2	AC027491 Homo sapi
23	22	22.0	71945	2	AC090754 Homo sapi
24	22	22.0	72784	2	AC100279 Mus muscu
25	22	22.0	83789	2	AC021679 Homo sapi
26	22	22.0	91956	2	AC021734 Homo sapi
27	22	22.0	98577	2	AC148154 Medicago
28	22	22.0	107287	2	AC146794 Medicago
29	22	22.0	110000	2	AC113552 Rattus no
30	22	22.0	110000	2	Continuation (3 of
31	22	22.0	110000	2	LMFLCHR28_0
32	22	22.0	115981	2	AL591072 Homo sapi
33	22	22.0	129503	2	AC044863 Mus muscu
34	22	22.0	138177	2	AC021931 Homo sapi
35	22	22.0	156661	2	AC148358 Homo sapi
36	22	22.0	159167	2	DMBR42L18
37	22	22.0	172742	2	AC023406 Homo sapi
38	22	22.0	176144	2	AC012408 Homo sapi
39	22	22.0	184050	2	AC087819 Homo sapi
40	22	22.0	188539	2	AC151100 Bos tauru
41	22	22.0	189053	2	AC016015 Homo sapi
42	22	22.0	192142	2	AC136061 Rattus no
43	22	22.0	193286	2	AC137285 Rattus no
44	22	22.0	199255	2	AC084195 Homo sapi
45	22	22.0	205662	2	AC145619 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS BD006932
DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION BD006932
VERSION BD006932.1 GI:18635303
KEYWORDS JP 2001503249-A/1.
SOURCE Lactococcus lactis
ORGANISM Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE 1 (bases 1 to 100)

AUTHORS Hammer, K. and Janssen, P.R.
TITLE Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL Patent: JP 2001503249-A 1 13-MAR-2001;

COMMENT PETER RUDAR JANSSEN

OS Lactococcus lactis

PN JP 2001503249-A/1

PD 13-MAR-2001

PF 25-AUG-1997 JP 1998510287

PR 23-AUG-1996 DK 0866/96

PI KALIN HAMMER, PETER RUDAR JANSSEN

CC C12N15/09, C12N15/00

CC Strandedness: Double;

CC Topology: Linear;

CC Key

FT promoter

FT -35signal

FT -10signal

FEATURES

Location/Qualifiers
1..100

/organism="Lactococcus lactis"

/mol_type="genomic DNA"

/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 100;

Best Local Similarity 100.0%; Pred. No. 2.6e-45;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	22	23	0	676	7	CV519050	0089P0022	
2	23	0	847	5	BQ58855			
3	22	22	0	403	9	CR400602	ArabIdops2	
4	22	22	0	430	6	CB798330		
5	21	21	0	339	9	CG988771	AMGNNTC:N	
6	21	21	0	389	9	CG988120	CB767001	
7	21	21	0	417	6	CB767001	AMGNNTC:S	
8	21	21	0	427	6	CB794660	AMGNNTC:S	
9	21	21	0	455	6	CB741719		
10	21	21	0	509	6	CB782841		
11	21	21	0	535	9	CG987802		
12	21	21	0	545	9	CG991922		
13	21	21	0	553	6	CB584453		
14	21	21	0	593	6	CB584453	AMGNNTC:S	
15	21	21	0	751	7	CN054920		
16	21	21	0	778	9	CG985602		
17	21	21	0	779	7	CN043556		
18	21	21	0	803	6	CA124432		
19	21	21	0	1063	7	W14754		
20	20	0	72	9	AL759144	ArabIdops		
21	20	20	0	83	9	AL759145	ArabIdops	
22	20	20	0	118	9	BX893640	ArabIdops	
23	20	20	0	153	9	BX534689	ArabIdops	
24	20	20	0	179	9	CL688872	ArabIdops	
	20	20	0	184	9	BX532547	ArabIdops	

c	25	20	20.0	184	9	EX948834	BX948834	ArabiDops
c	26	20	20.0	186	9	EX293365	BX293365	ArabiDops
c	27	20	20.0	203	9	EX894565	BX894565	ArabiDops
c	28	20	20.0	213	9	CR357873	CR357873	ArabiDops
c	29	20	20.0	224	9	CR400397	CR400397	ArabiDops
c	30	20	20.0	242	9	EX661521	EX661521	ArabiDops
c	31	20	20.0	247	9	CR400564	CR400564	ArabiDops
c	32	20	20.0	250	9	CR396322	CR396322	ArabiDops
c	33	20	20.0	255	9	CR402333	CR402333	ArabiDops
c	34	20	20.0	274	9	EX891296	EX891296	ArabiDops
c	35	20	20.0	278	9	CG977426	CG977426	CH240.167
c	36	20	20.0	299	6	CB701255	CB701255	AMGNNTC:M
c	37	20	20.0	304	9	EX945052	EX945052	ArabiDops
c	38	20	20.0	307	9	EX895939	EX895939	ArabiDops
c	39	20	20.0	328	9	CG561371	CG561371	OST182853
c	40	20	20.0	332	9	EX292036	BX292036	ArabiDops
c	41	20	20.0	335	9	CR358353	CR358353	ArabiDops
c	42	20	20.0	336	9	CG977455	CG977455	CH240.167
c	43	20	20.0	352	9	EX948732	EX948732	ArabiDops
c	44	20	20.0	353	9	CL607842	CL607842	CH240.174
c	45	20	20.0	354	9	EX291408	BX291408	ArabiDops

ALIGNMENTS

RESULT 1	
CV519050	
LOCUS	676 bp mRNA linear EST 06-OCT-2004
DEFINITION	0089P00222.x0_H04 Mimulus guttatus library 2 Mimulus guttatus cDNA clone 0089P00222.x0_H04, mRNA sequence.
ACCESSION	CV519050
VERSION	
KEYWORDS	
SOURCE	CV519050.1 GI:53845582
ORGANISM	Mimulus guttatus (spotted monkey flower)
	Mimulus guttatus
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Lamiales; Lamiales incertae sedis; Mimosaceae; Mimulus.
REFERENCE	1. (bases 1 to 676)
AUTHORS	Willis,J., Vision,T., Dietrich,F.S. and Allen,A.
TITLE	Mimulus guttatus cDNA sequence
JOURNAL	Unpublished (2004)
COMMENT	Contact: Willis J Department of Biology Duke University 072-A Biological Sciences Science Drive, Durham, NC 27708, USA Tel.: 919 660 7340 Fax: 919 660 7293 Email: jwillis@duke.edu Plate: 0089P0022 row: 04 column: H Seq primer: T7 High quality sequence start: 17 High quality sequence stop: 514. Location/Qualifiers 1..676 /organism="Mimulus guttatus" /mol_type="mRNA" /db_xref="taxon:4155" /clone="0089P00222.x0_H04" /note="Vector: pGEM-T Easy; a Mimulus guttatus cDNA library"
FEATURES	
Source	
ORIGIN	
Query Match	23.0%; Score 23; DB 7; Length 676;
Best Local Similarity	100.0%; Pred. No. 0.032;
Matches	23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
cy	37 TTCTTGCACNNNNNNNNNNNNN 59
cb	37 TTCTTGCACNNNNNNNNNNNNN 59

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 223.488 Seconds

(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-1

Perfect score: 100

Sequence: 1 CGGGATCCTTAAGATATTA.....TTACTGACGTGAATTCGG 100

Scoring table: OLIGO_NUC

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseq1980s:*

2: geneseq1990s:*

3: geneseq2000s:*

4: geneseq2001as:*

5: geneseq2002as:*

6: geneseq2002bs:*

7: geneseq2003as:*

8: geneseq2003bs:*

9: geneseq2003cs:*

10: geneseq2003ds:*

11: geneseq2004as:*

12: geneseq2004bs:*

13: geneseq2004ds:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	100	2	AAV23173
2	41	41.0	113	2	AAV23174
3	24	24.0	147443	11	ACMA4752
4	22	22.0	1384	5	ABV25481
5	22	22.0	263744	10	ADFO8271
6	21	21.0	500	12	ADL86200
7	21	21.0	500	12	ADL86201
8	20	20.0	28	2	AAV32732
9	20	20.0	28	2	AAQ85563
10	20	20.0	28	2	AAQ85562
11	20	20.0	28	2	AAV09634
12	20	20.0	28	2	AAV32733
13	20	20.0	30	2	AAV32734
14	20	20.0	30	8	ABX14327
15	20	20.0	43	6	ABX52819
16	20	20.0	52	12	ADQ57897
17	20	20.0	67	8	ABZ27733
18	20	20.0	73	8	ABZ79664
19	20	20.0	336	2	AAQ70086
20	20	20.0	418	12	ADL84156

c	21	20	20.0	418	12	ADL84157	ADL84157 DNA up-re
c	22	20	20.0	674	6	ABK53696	ABK53696 Human eos
c	23	20	20.0	1999	10	ADCO8481	ADCO8481 Rice DNA
c	24	20	20.0	2763	2	AAV15167	AAV15167 Methylthi
c	25	20	20.0	2763	2	AAQ92813	AAQ92813 Human MTA
c	26	20	20.0	2763	2	AAV85305	AAV85305 Human met
c	27	20	20.0	2763	2	AAZ00866	AAZ00866 Human MTA
c	28	20	20.0	2763	5	AAV86091	AAV86091 Methylthi
c	29	20	20.0	2763	10	AAV64097	AAV64097 Human met
c	30	20	20.0	2784	2	AAQ9202	AAQ9202 Pseudomon
c	31	20	20.0	2860	12	ADQ23151	ADQ23151 Human sof
c	32	20	20.0	5117	12	ADQ24050	ADQ24050 Human sof
c	33	20	20.0	9179	2	AAV74411	AAV74411 Staphyloc
c	34	20	20.0	21808	13	ADV67055	ADV67055 Mouse can
c	35	20	20.0	59590	3	AAV2281	AAV2281 BAC conta
c	36	20	20.0	103661	12	ADQ97712	ADQ97712 Mouse can
c	37	20	20.0	110000	13	ABD32594	ABD32594 (3 of
c	38	20	20.0	11836	13	ABD33102	ABD33102 Murine ca
c	39	20	20.0	215248	12	ADQ97284	ADQ97284 Mouse can
c	40	19	19.0	100	10	ADB87585	ADB87585 Primer DN
c	41	19	19.0	106	2	AAQ48612	AAQ48612 Synthetic
c	42	19	19.0	106	2	ADL88439	ADL88439 DNA up-re
c	43	19	19.0	290	12	ADL88438	ADL88438 DNA up-re
c	44	19	19.0	321	12	ADQ80615	ADQ80615 Human lig
c	45	19	19.0				

ALIGNMENTS

RESULT 1	AAV23173	standard; DNA; 100 BP.
ID	AAV23173	
XX	AAV23173;	
XX	28-UTL-1998	(first entry)
XX	Lactococcus lactis	consensus promoter sequence.
XX	Lactococcus lactis	promoter; optimise; gene expression; spacer;
KM	artificial promoter library; de.	
XX	Synthetic.	
OS	Lactococcus lactis.	
XX	Key	Location/Qualifiers
FT	promoter	/tag= a
FT		/note= "artificial promoter library"
FT	misc_feature	/tag= b
FT		/note= "consensus sequence"
FT	-35_signal	/tag= c
FT		/tag= d
FT	misc_feature	/tag= e
FT		/tag= f
FT	-10_signal	/tag= g
FT		/tag= h
FT	misc_feature	/tag= i
FT		/tag= j
FT		/tag= k
FT		/tag= l
FT		/tag= m
FT		/tag= n
FT		/tag= o
FT		/tag= p
FT		/tag= q
FT		/tag= r
FT		/tag= s
FT		/tag= t
FT		/tag= u
FT		/tag= v
FT		/tag= w
FT		/tag= x
FT		/tag= y
FT		/tag= z
FT		/tag= AA
FT		/tag= AB
FT		/tag= AC
FT		/tag= AD
FT		/tag= AE
FT		/tag= AF
FT		/tag= AG
FT		/tag= AH
FT		/tag= AI
FT		/tag= AJ
FT		/tag= AK
FT		/tag= AL
FT		/tag= AM
FT		/tag= AN
FT		/tag= AO
FT		/tag= AP
FT		/tag= AQ
FT		/tag= AR
FT		/tag= AS
FT		/tag= AT
FT		/tag= AU
FT		/tag= AV
FT		/tag= AW
FT		/tag= AX
FT		/tag= AY
FT		/tag= AZ
FT		/tag= BA
FT		/tag= BB
FT		/tag= BC
FT		/tag= BD
FT		/tag= BE
FT		/tag= BF
FT		/tag= BG
FT		/tag= BH
FT		/tag= BI
FT		/tag= BJ
FT		/tag= BK
FT		/tag= BL
FT		/tag= BM
FT		/tag= BN
FT		/tag= BO
FT		/tag= BP
FT		/tag= BQ
FT		/tag= BR
FT		/tag= BS
FT		/tag= BT
FT		/tag= BU
FT		/tag= BV
FT		/tag= BW
FT		/tag= BX
FT		/tag= BY
FT		/tag= BZ
FT		/tag= CA
FT		/tag= CB
FT		/tag= CC
FT		/tag= CD
FT		/tag= CE
FT		/tag= CF
FT		/tag= CG
FT		/tag= CH
FT		/tag= CI
FT		/tag= CJ
FT		/tag= CK
FT		/tag= CL
FT		/tag= CM
FT		/tag= CN
FT		/tag= CO
FT		/tag= CP
FT		/tag= CQ
FT		/tag= CR
FT		/tag= CS
FT		/tag= CT
FT		/tag= CU
FT		/tag= CV
FT		/tag= CW
FT		/tag= CX
FT		/tag= CY
FT		/tag= CZ
FT		/tag= DA
FT		/tag= DB
FT		/tag= DC
FT		/tag= DD
FT		/tag= DE
FT		/tag= DF
FT		/tag= DG
FT		/tag= DH
FT		/tag= DI
FT		/tag= DJ
FT		/tag= DK
FT		/tag= DL
FT		/tag= DM
FT		/tag= DN
FT		/tag= DO
FT		/tag= DP
FT		/tag= DQ
FT		/tag= DR
FT		/tag= DS
FT		/tag= DT
FT		/tag= DU
FT		/tag= DV
FT		/tag= DW
FT		/tag= DX
FT		/tag= DY
FT		/tag= DZ
FT		/tag= EA
FT		/tag= EB
FT		/tag= EC
FT		/tag= ED
FT		/tag= EE
FT		/tag= EF
FT		/tag= EG
FT		/tag= EH
FT		/tag= EI
FT		/tag= EJ
FT		/tag= EK
FT		/tag= EL
FT		/tag= EM
FT		/tag= EN
FT		/tag= EO
FT		/tag= EP
FT		/tag= EQ
FT		/tag= ER
FT		/tag= ES
FT		/tag= ET
FT		/tag= EU
FT		/tag= EV
FT		/tag= EW
FT		/tag= EX
FT		/tag= EY
FT		/tag= EZ
FT		/tag= FA
FT		/tag= FB
FT		/tag= FC
FT		/tag= FD
FT		/tag= FE
FT		/tag= FF
FT		/tag= FG
FT		/tag= FH
FT		/tag= FI
FT		/tag= FJ
FT		/tag= FK
FT		/tag= FL
FT		/tag= FM
FT		/tag= FN
FT		/tag= FO
FT		/tag= FP
FT		/tag= FQ
FT		/tag= FR
FT		/tag= FS
FT		/tag= FT
FT		/tag= FU
FT		/tag= FV
FT		/tag= FW
FT		/tag= FX
FT		/tag= FY
FT		/tag= FZ
FT		/tag= GA
FT		/tag= GB
FT		/tag= GC
FT		/tag= GD
FT		/tag= GE
FT		/tag= GF
FT		/tag= GG
FT		/tag= GH
FT		/tag= GI
FT		/tag= GJ
FT		/tag= GK
FT		/tag= GL
FT		/tag= GM
FT		/tag= GN
FT		/tag= GO
FT		/tag= GP
FT		/tag= GQ
FT		/tag= GR
FT		/tag= GS
FT		/tag= GT
FT		/tag= GU
FT		/tag= GV
FT		/tag= GW
FT		/tag= GX
FT		/tag= GY
FT		/tag= GZ
FT		/tag= HA
FT		/tag= HB
FT		/tag= HC
FT		/tag= HD
FT		/tag= HE
FT		/tag= HF
FT		/tag= HG
FT		/tag= HH
FT		/tag= HI
FT		/tag= HJ
FT		/tag= HK
FT		/tag= HL
FT		/tag= HM
FT		/tag= HN
FT		/tag= HO
FT		/tag= HP
FT		/tag= HQ
FT		/tag= HR
FT		/tag= HS
FT		/tag= HT
FT		/tag= HU
FT		/tag= HV
FT		/tag= HW
FT		/tag= HX
FT		/tag= HY
FT		/tag= HZ
FT		/tag= IA
FT		/tag= IB
FT		/tag= IC
FT		/tag= ID
FT		/tag= IE
FT		/tag= IF
FT		/tag= IG
FT		/tag= IH
FT		/tag= II
FT		/tag= IJ
FT		/tag= IK
FT		/tag= IL
FT		/tag= IM
FT		/tag= IN
FT		/tag= IO
FT		/tag= IP
FT		/tag= IQ
FT		/tag= IR
FT		/tag= IS
FT		/tag= IT
FT		/tag= IU
FT		/tag= IV
FT		/tag= IW
FT		/tag= IX
FT		/tag= IY
FT		/tag= IZ
FT		/tag= JA
FT		/tag= JB
FT		/tag= JC
FT		/tag= JD
FT		/tag= JE
FT		/tag= JF
FT		/tag= JG
FT		/tag= JH
FT		/tag= JI
FT		/tag= JJ
FT		/tag= JK
FT		/tag= JL
FT		/tag= JM
FT		/tag= JN
FT		/tag= JO
FT		/tag= JP
FT		/tag= JQ
FT		/tag= JR
FT		/tag= JS
FT		/tag= JT
FT		/tag= JU
FT		/tag= JV
FT		/tag= JW
FT		/tag= JX
FT		/tag= JY
FT		/tag= JZ
FT		/tag= KA
FT		/tag= KB
FT		/tag= KC
FT		/tag= KD
FT		/tag= KE
FT		/tag= KF
FT		/tag= KG
FT		/tag= KH
FT		/tag= KI
FT		/tag= KJ
FT		/tag= KK
FT		/tag= KL

Result No.	Score	Query Match	Length	DB	ID	Description
1	113	100.0	113	6	BD006933	BD006933 Artificia
2	41	36.3	100	6	BD006932	BD006932 Artificia
3	31	27.4	113	6	BD006933	BD006933 Artificia
4	27	23.9	27	6	AX839866	AX839866 Sequence
5	27	23.9	27	6	AX839866	AX839866 Sequence
6	26	23.0	215551	2	AC127000	AC127000 Rattus no
7	25	22.1	71166	2	AC023783	AC023783 Homo sapi
8	25	22.1	157382	2	AC009279	AC009279 Homo sapi
9	25	22.1	159403	2	AC024735	AC024735 Homo sapi
10	25	22.1	159560	2	AC028955	AC028955 Homo sapi
11	25	22.1	192713	2	AC116253	AC116253 Rattus no
12	24	21.2	55730	2	AC102301	AC102301 Mus muscu
13	24	21.2	72487	2	AC115951	AC115951 Mus muscu
14	24	21.2	104757	2	AF000625	AF000625 Homo sapi
15	24	21.2	110000	2	AC091367	Continuation (5 oclan
16	24	21.2	128986	2	AC143183	AC143183 Macaca mu
17	24	21.2	145986	2	AC022467	AC022467 Homo sapi
18	24	21.2	225784	2	AC015965	AC015965 Homo sapi
19	23	20.4	53949	2	AC090724	AC090724 Homo sapi

C	20	23	20.4	63125	2	AC134783	AC134783 Homo sapi
C	21	23	20.4	74105	2	AC087648	AC087648 Homo sapi
C	22	23	20.4	110000	2	AC118330 ⁻⁰	AC118330 Rattus no
C	23	23	20.4	110000	2	AC118330 ⁻¹	Continuation (2 of
C	24	23	20.4	117445	2	AL160263	AL160263 Homo sapi
C	25	23	20.4	120444	2	AC136092	AC136092 Rattus no
C	26	23	20.4	164840	2	AC011127	AC011127 Homo sapi
C	27	23	20.4	175654	2	AC025496	AC025496 Homo sapi
C	28	23	20.4	189556	2	AC141170	AC141170 Rattus no
C	29	23	20.4	212722	2	AC087158	AC087158 Mus muscu
C	30	23	20.4	214258	2	AC132994	AC132994 Rattus no
C	31	23	20.4	220434	2	AC106486	AC106486 Rattus no
C	32	23	20.4	227070	2	AC103323	AC103323 Rattus no
C	33	23	20.4	228108	2	AC142063	AC142063 Rattus no
C	34	23	20.4	232993	2	AC098761	AC098761 Rattus no
C	35	23	20.4	238883	2	AC122107	AC122107 Rattus no
C	36	23	20.4	241416	2	AC114100	AC114100 Rattus no
C	37	23	20.4	247062	2	AC099455	AC099455 Rattus no
C	38	23	20.4	292946	2	AC099347	AC099347 Rattus no
C	39	23	20.4	295808	2	AC151867	AC151867 Callithri
C	40	22	19.5	1384	6	C0493603	C0493603 Sequence
C	41	22	19.5	37661	2	AC150980	AC150980 Medicago
C	42	22	19.5	49070	2	AC101123	AC101123 Mus muscu
C	43	22	19.5	55208	2	AC087394	AC087394 Homo sapi
C	44	22	19.5	59935	2	AC018376	AC018376 Homo sapi
C	45	22	19.5	61258	2	AC023082	AC023082 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BD006933	113 bp	Artificial promoter libraries for selected organisms and promoters	BD006933	BD006933.1	GI:18635304	JP 2001503249-A/2.	unidentified unidentified unclassified.	1 (bases 1 to 113) Hammer K. and Janssen P.R. Artificial promoter libraries for selected organisms and promoters derived from such libraries Patent: JP 2001503249-A 2 13-MAR-2001;	PETER RUDAR JANSSEN	OS	Unidentified	

FEATURES

```
/organism="unidentified"  
/mol_type="genomic DNA"  
/db_xref="taxon:32644"
```

ORIGIN

Query Match	100.0%	Score 113;	DB 6;	Length 113;
Best Local Similarity	100.0%;	Pred. No. 2.6e-52;		
Matches 113; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 CGGATCCAAGCTTATATTAATTAGCACTCNNNNNNNNNGAGTGCTAATTTTTGACA 60

Result	Query	Score	No.	ID	Description
C 1	23	20.4	435	6	CB546770 AMANNUC:1
C 2	23	20.4	438	6	CB790308 AMANNUC:1
C 3	23	20.4	568	7	CF918345 Bfior498
C 4	23	20.4	808	6	CB898886 tr1c016xtr
5	23	20.4	847	5	BQ589855 E012562-C
6	22	19.5	299	6	CB701255
7	22	19.5	390	6	CB778306
8	22	19.5	525	5	CG987802
C 9	22	19.5	810	9	CC920506
C 10	22	19.5	839	7	CR287310
C 11	22	19.5	9155	9	AY417491
C 12	21	18.6	184	9	BX948834
C 13	21	18.6	339	9	CG988771
C 14	21	18.6	389	9	CG988120
C 15	21	18.6	435	6	CB792605
C 16	21	18.6	447	6	CB746942
17	21	18.6	514	9	CG547018
18	21	18.6	545	9	CG919192
19	21	18.6	751	7	CN054920
C 20	21	18.6	779	7	CN043552
C 21	21	18.6	815	9	CL668316
C 22	21	18.6	866	7	CK157343
C 23	21	18.6	1419	9	CK167285
C 24	21	18.6	9166	9	AY417490

C	25	21	18.6	9177	9	AAY17489
C	26	20	17.7	722	9	AL759144
C	27	20	17.7	83	9	AL759145
C	28	20	17.7	118	9	BX8935c40
C	29	20	17.7	118	9	BX8935c40
C	30	20	17.7	179	9	BX534689
C	31	20	17.7	184	9	CL688872
C	32	20	17.7	184	9	BX532547
C	33	20	17.7	186	9	BX292365
C	34	20	17.7	203	9	BX894565
C	35	20	17.7	213	9	CR357873
C	36	20	17.7	224	9	CR400397
C	37	20	17.7	242	9	BX661521
C	38	20	17.7	247	9	CR400564
C	39	20	17.7	250	9	CR396322
C	40	20	17.7	255	9	CR402333
C	41	20	17.7	274	9	BX891296
C	42	20	17.7	278	9	CG977426
C	43	20	17.7	304	9	BX945052
C	44	20	17.7	307	9	BX895839
C	45	20	17.7	328	9	CG561371
C	46	20	17.7	332	9	BX292036

ALIGNMENTS

[illegible]

09 NNNNNNNNNNNNNNNNNNNNNNNN 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 444.741 Seconds

2648.796 Million cell updates/sec

Title: US-09-242-657B-3
Perfect score: 100

Sequence: 1 CAGAAATTCGTACTCANNNN.....GCTACCAATCATGATCCCG 199

Scoring table: OLIGO_NUC

Searched: 4390206 seqs, 2959870667 residues

word size : 0

Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
Maximum DB seq length: 3000000000
```

Maximum DB seq length: 200000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

```

1: N_genseq_16dec04:*
2: genseq_1980s:*
3: genseq_190s:*
4: genseq_2000s:*
5: genseq_2001as:*
6: genseq_2001bs:*
7: genseq_2002as:*
8: genseq_2002bs:*
9: genseq_2003as:*
10: genseq_2003bs:*
11: genseq_2003cs:*
12: genseq_2004as:*
13: genseq_2004bs:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	199	100.0	199	2	AAV33176	AAV33176 Saccharom
2	76	38.2	167	2	AAV23218	AAV23218 Saccharom
3	76	38.2	177	2	AAV33215	AAV33215 Saccharom
4	76	38.2	195	2	AAV33225	AAV33225 Saccharom
5	76	38.2	195	2	AAV33227	AAV33227 Saccharom
6	76	38.2	195	2	AAV33220	AAV33220 Saccharom
7	68	34.2	189	2	AAV33228	AAV33228 Saccharom
8	68	34.2	195	2	AAV33222	AAV33222 Saccharom
9	67	33.7	3720	6	ABK87221	ABK87221 Human lip
c 10	66	33.2	101	4	AAK50304	AAK50304 RSO1 sequ
c 11	66	33.2	375	2	AAV78357	AAV78357 Staphyloc
12	66	33.2	400	2	AAV78167	AAV78167 Staphyloc
13	66	33.2	469	2	AAV75210	AAV75210 Staphyloc
14	66	33.2	571	2	AAV75287	AAV75287 Staphyloc
15	66	33.2	782	2	AAV75008	AAV75008 Staphyloc
16	66	33.2	919	2	AAV74871	AAV74871 Staphyloc
c 17	66	33.2	2881	2	AAV74565	AAV74565 Staphyloc
18	66	33.2	4253	2	AAV74428	AAV74428 Staphyloc
19	66	33.2	10182	2	AAV74427	AAV74427 Staphyloc
c 20	66	33.2	11802	2	AAV74381	AAV74381 Staphyloc

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-5

Perfect score: 60
Sequence: 1 CATACCGAGTTATCTCTG.....TATATATCTCAGTACTGT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBml:.*
1: gb ba:.*
2: gb_btg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sbs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006936	BD006936 Artificialia
2	20	33.3	60	BD006950	BD006950 Artificialia
3	19	31.7	228554	AC108956	AC108956 Rattus no
4	19	31.7	238316	AC094974	AC094974 Rattus no
5	18	30.0	60	BD006941	BD006941 Artificialia
6	18	30.0	60	BD006948	BD006948 Artificialia
7	18	30.0	60	BD006957	BD006957 Artificialia
8	18	30.0	2482	AK094737	AK094737 Homo sapi
9	18	30.0	23326	LC012427	LC012427 Homo sapi
10	18	30.0	115928	AC127427	AC127427 Magnapor
11	18	30.0	128379	HSJ773A18	HSJ773A18 Human DNA
12	18	30.0	129532	AP004482	AP004482 Lotus cor
13	18	30.0	137416	AC092705	AC092705 Homo sapi
14	18	30.0	137593	AP001406	AP001406 Homo sapi
15	18	30.0	156766	AP002400	AP002400 Homo sapi
16	18	30.0	157175	AC011011	AC011011 Homo sapi
17	18	30.0	151443	AC009884	AC009884 Homo sapi
18	18	30.0	157949	AC087821	AC087821 Homo sapi
19	18	30.0	168838	AC024929	AC024929 Homo sapi

C	20	18	30.0	174278	2	AC068344	AC068344 Homo sapi
C	21	18	30.0	177999	2	AC068154	AC068154 Homo sapi
C	22	18	30.0	183044	2	AC079827	AC079827 Homo sapi
C	23	18	30.0	187579	2	AP001450	AP001450 Homo sapi
C	24	18	30.0	194833	2	AP001452	AP001452 Homo sapi
C	25	17	28.3	59	6	BD006969	BD006969 Artificialia
C	26	17	28.3	60	6	BD006958	BD006958 Artificialia
C	27	17	28.3	60	6	BD006959	BD006959 Artificialia
C	28	17	28.3	60	6	BD006963	BD006963 Artificialia
C	29	17	28.3	60	6	BD006970	BD006970 Artificialia
C	30	17	28.3	1871	8	AK101907	AK101907 Oryza sat
C	31	17	28.3	14417	3	U28409	U28409 Caenorhabdi
C	32	17	28.3	47163	10	AL954297	AL954297 Mouse DNA
C	33	17	28.3	73148	2	AC048349	AC048349 Homo sapi
C	34	17	28.3	73170	9	AL445068	AL445068 Human DNA
C	35	17	28.3	84076	2	AC149797	AC149797 Rhes aeg
C	36	17	28.3	92501	8	CR380956	CR380956 (12 o
C	37	17	28.3	93003	9	AF547386	AF547386 Homo sapi
C	38	17	28.3	107315	9	AL136303	AL136303 Human DNA
C	39	17	28.3	122538	9	AC011942	AC011942 Homo sapi
C	40	17	28.3	122939	8	AC116369	AC116369 Oryza sat
C	41	17	28.3	145729	9	AC117489	AC117489 Homo sapi
C	42	17	28.3	153964	9	AL139377	AL139377 Human DNA
C	43	17	28.3	155744	2	AL157784	AL157784 Homo sapi
C	44	17	28.3	156840	9	AC006036	AC006036 Homo sapi
C	45	17	28.3	160669	2	AC104477	AC104477 Pan trogl

ALIGNMENTS

RESULT 1
LOCUS BD006936
DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION BD006936.1 GI:18635307
VERSION JP 2001503249-A/5.
KEYWORDS Lactococcus Lactis
SOURCE Lactococcus Lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE

AUTHORS Hammer,K. and Janssen,P.R.
TITLE Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL Patent: JP 2001503249-A 5 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Lactococcus Lactis
PN JP 2001503249-A/5
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09,C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (4)..(60).
Location/Qualifiers
1..60
/organism="Lactococcus Lactis"
/mol_type="Genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATACCGAGTTATCTTGACAGTTCACCTCGGGTGTATATATCTCAGTACTGT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-5

Perfect score: 60

Sequence: 1 CATACCGAGTTTATCTCTG.....TATAATATCTAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database:

N.Geneseq.16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23177 Lactococc
2	20	33.3	60	2	AAV23191 Lactococc
3	18	30.0	60	2	AAV23182 Lactococc
4	18	30.0	60	2	AAV23189 Lactococc
5	18	30.0	60	2	AAV23198 Lactococc
6	18	25392	4	AAK82159	AAK82159 Human Imm
7	18	30.0	41787	13	ABD33599 Human can
8	17	28.3	59	2	AAV23310 Lactococc
9	17	28.3	60	2	AAV23311 Lactococc
10	17	28.3	60	2	AAV23300 Lactococc
11	17	28.3	60	2	AAV23204 Lactococc
12	17	28.3	60	2	AAV23199 Lactococc
13	16	26.7	58	2	AAV23184 Lactococc
14	16	26.7	58	6	ABK98625 L. lactis
15	16	26.7	58	2	ABK98625 L. lactis
16	16	26.7	59	2	AAV23190 Lactococc
17	16	26.7	59	2	AAV23185 Lactococc
18	16	26.7	60	2	AAV23203 Lactococc
19	16	26.7	60	2	AAV23207 Lactococc
20	16	26.7	60	2	AAV23178 Lactococc

ALIGNMENTS

21	16	26.7	64	6	ABK98604	ABK98604 L. lactis
22	16	26.7	64	9	ACD13855	ACD13855 L. lactis
23	16	26.7	65	9	ABK98605	ABK98605 L. lactis
24	16	26.7	65	9	ACD13856	ACD13856 L. lactis
25	16	26.7	93	6	ABK98615	ABK98615 L. lactis
26	16	26.7	93	9	ACD13866	ACD13866 L. lactis
27	16	26.7	265	2	AAO23880	AAO23880 pKTH1820
28	16	26.7	265	2	AAO46205	AAO46205 Promoter
29	16	26.7	265	2	AAT31875	AAT31875 Promoter
30	16	26.7	283	9	ADAS9739	ADAS9739 Soybean f
31	16	26.7	283	12	ADP61080	ADP61080 Soybean c
32	16	26.7	296	3	AAC03364	AAC03364 Human sec
33	16	26.7	705	12	ADO03462	ADO03462 Thalecres
34	16	26.7	755	6	ABN99221	ABN99221 Arabidops
35	16	26.7	983	6	ABT07262	ABT07262 Human Cpg
36	16	26.7	1062	6	ABZ13652	ABZ13652 Arabidops
37	16	26.7	1062	6	ADG87654	ADG87654 A. thalia
38	16	26.7	1062	6	ADG87655	ADG87655 A. thalia
39	16	26.7	1062	8	ADA68053	ADA68053 Arabidops
40	16	26.7	1305	3	AAC47953	AAC47953 Arabidops
41	16	26.7	1305	3	AAC51562	AAC51562 Arabidops
42	16	26.7	2118	3	AAC51464	AAC51464 Arabidops
43	16	26.7	2707	10	ADD71194	ADD71194 Human int
44	16	26.7	2864	12	ADP50636	ADP50636 Marine DN
45	16	26.7	9417	6	ABK98632	ABK98632 Vector pE

RESULT 1

AAV23177 standard; DNA; 60 BP.

AAV23177;

28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cpl.

Lactococcus lactis; constitutional promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Synthetic.

Lactococcus lactis.

Key

promoter

4. .60

Location/Qualifiers

/*tag= a

/standard_name= "Constitutional promoter"

WO9807846-A1.

26-FEB-1998.

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENS/) JENSEN P. R.

Jensen PR, Hammer K;

WPI; 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 43; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-5

Perfect score: 60
Sequence: 1 CATACCGAGTTTATCTCTG.....TATTAATATCAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	31.7	692	6	CA072590 SCEPAM101
C 2	19	31.7	912	7	CA072590 SCEPAM101
C 3	18	30.0	161	9	CE525654 tigr-988
C 4	18	30.0	407	8	BH506765 BOHOC87TF
C 5	18	30.0	506	6	CB119387 K-EST0166
C 6	18	30.0	613	9	CE039872 tigr-988
C 7	18	30.0	630	5	BK473371 DKEZp886U
C 8	18	30.0	706	4	BI088532 602853061
C 9	18	30.0	927	5	BU189979 AGENCOURT
C 10	18	30.0	1116	4	BI910285 603068325
C 11	17	28.3	416	9	CE287823 tigr-988
C 12	17	28.3	480	1	AU210653 tigr-988
C 13	17	28.3	500	6	CA995454 tigr-988
C 14	17	28.3	581	7	CF426773 tigr-988
C 15	17	28.3	581	7	CF426773 tigr-988
C 16	17	28.3	583	8	BZ295332 CG1425.t1
C 17	17	28.3	603	8	BE145823 MRO-HT020
C 18	17	28.3	628	7	CF416219 tigr-988
C 19	17	28.3	651	4	BG866487 602785436
C 20	17	28.3	671	6	CA159481 SCFR23CO
C 21	17	28.3	674	6	CA159393 SCFR23CO
C 22	17	28.3	680	9	CE018699 tigr-988
C 23	17	28.3	687	9	CL746751 OR_BBA011
C 24	17	28.3	699	8	BZ028812 oeg98604

C 25	17	28.3	711	7	CV259956 MS02012.B
C 26	17	28.3	725	9	CL789447 OR_BBA010
C 27	17	28.3	771	8	BH532824 BOGLP79TR
C 28	17	28.3	786	8	BH579089 BOGNC57P
C 29	17	28.3	790	9	CC863504 ND1.41P16
C 30	17	28.3	812	9	CL611981 OR_BBA000
C 31	17	28.3	844	9	CG953853 MBEDK10TF
C 32	17	28.3	847	8	BH605925 BOHUR80TR
C 33	17	28.3	856	4	BF972618 602243049
C 34	17	28.3	865	6	CA766325 AF53-RP6
C 35	17	28.3	927	9	CG703804 ZMMBB0011
C 36	17	28.3	927	9	CG997598 ZMMBB005
C 37	17	28.3	971	9	CL063262 CH216-980
C 38	16	26.7	101	7	CV889407 QV2-BT068
C 39	16	26.7	171	2	BF332389 QV2-BT068
C 40	16	26.7	180	1	AV181251 AV181251
C 41	16	26.7	206	1	AI600637 486060D07
C 42	16	26.7	217	9	CR046864 Reverse 8
C 43	16	26.7	231	9	AG251477 locus cor
C 44	16	26.7	259	2	BF332375 QV2-BT068
C 45	16	26.7	270	2	BB408680 BB408680

ALIGNMENTS

RESULT 1
CA072590 692 bp mRNA linear EST 23-SRP-2003
LOCUS SCEPAM1018A05.g AM1 Saccharum officinarum cDNA clone SCEPAM1018A05
DEFINITION 5' mRNA sequence.
CA072590
CA072590.1 GI:34924741
VERSION
KEYWORDS
SOURCE
ORGANISM
Saccharum officinarum
Saccharum officinarum
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;
Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, PACCAD
clade, Panicoideae, Andropogoneae, Saccharum, Saccharum officinarum
complex.
1 (bases 1 to 692)
REFERENCE
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
JOURNAL
COMMENT
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: pattunda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 018 row: A column: 05
Seq primer: TV Promoter Primer.
Location:Qualifiers
1..692
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCEPAM1018A05"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site 1: Salt; Site 2:
Nori. An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a agarose
CL-2B 40cm-columns and fragments sitting between 0.8 and
1.5 Kb were directionally cloned into the vector. Details

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-6

Perfect score: 60

Sequence: 1 CANGGCTAGTTATCTCTG.....TATATAGGACAGTACTGTT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Genembl:*

1: gb_ba:*\n2: gb_btg:*\n3: gb_in:*\n4: gb_om:*\n5: gb_ov:*\n6: gb_pac:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_xo:*\n11: gb_ses:*\n12: gb_sy:*\n13: gb_un:*\n14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006937	Artificial promoter libraries for selected organisms and promoters
2	20	33.3	290386	AC094511	Artificial promoter libraries for selected organisms and promoters
3	19	31.7	525	BV184768	Artificial promoter libraries for selected organisms and promoters
4	19	31.7	96153	AC092651	Artificial promoter libraries for selected organisms and promoters
5	19	31.7	193136	AC151023	Artificial promoter libraries for selected organisms and promoters
6	19	31.7	209161	AC010913	Artificial promoter libraries for selected organisms and promoters
7	19	31.7	209980	AC151038	Artificial promoter libraries for selected organisms and promoters
8	18	30.0	34475	AC138692	Artificial promoter libraries for selected organisms and promoters
9	18	30.0	55801	AC087476	Artificial promoter libraries for selected organisms and promoters
10	18	30.0	69146	AC086668	Artificial promoter libraries for selected organisms and promoters
11	18	30.0	108885	AC023135	Artificial promoter libraries for selected organisms and promoters
12	18	30.0	133337	AC016231	Artificial promoter libraries for selected organisms and promoters
13	18	30.0	138404	AC117187	Artificial promoter libraries for selected organisms and promoters
14	18	30.0	146310	AC090459	Artificial promoter libraries for selected organisms and promoters
15	18	30.0	154513	AC145047	Artificial promoter libraries for selected organisms and promoters
16	18	30.0	158759	AC144916	Artificial promoter libraries for selected organisms and promoters
17	18	30.0	164290	AC022709	Artificial promoter libraries for selected organisms and promoters
18	18	30.0	167416	AC025856	Artificial promoter libraries for selected organisms and promoters
19	18	30.0	172149	BX546448	Human DNA

c	20	18	30.0	172810	9	AC145616	AC145616 Homo sapi
c	21	18	30.0	175667	2	AC069520	AC069520 Homo sapi
c	22	18	30.0	176368	2	AC140859	AC140859 Homo sapi
c	23	18	30.0	178925	9	AC145613	AC145613 Homo sapi
c	24	18	30.0	179138	2	AC121439	AC121439 Rattus no
c	25	18	30.0	180689	2	AC124913	AC124913 Sus scrof
c	26	18	30.0	182416	2	AC023044	AC023044 Homo sapi
c	27	18	30.0	182529	2	AC064832	AC064832 Homo sapi
c	28	18	30.0	184090	10	AC124197	AC124197 Mus muscu
c	29	18	30.0	184355	9	CR381572	CR381572 Human DNA
c	30	18	30.0	185637	9	AC145547	AC145547 Homo sapi
c	31	18	30.0	186012	9	AC068349	AC068349 Homo sapi
c	32	18	30.0	188992	9	AC147648	AC147648 Homo sapi
c	33	18	30.0	193321	10	AC121898	AC121898 Mus muscu
c	34	18	30.0	220303	2	AC111672	AC111672 Rattus no
c	35	18	30.0	234520	2	AC115233	AC115233 Rattus no
c	36	18	30.0	243969	5	AC145956	AC145956 Gallus ga
c	37	18	30.0	246443	2	AC112892	AC112892 Rattus no
c	38	18	30.0	262782	2	AC115230	AC115230 Rattus no
c	39	17	28.3	58	6	BD006943	BD006943 Artificia
c	40	17	28.3	60	6	BD006864	BD006864 Artificia
c	41	17	28.3	60	6	BD006972	BD006972 Artificia
c	42	17	28.3	2612	8	AF349514	AF349514 Aspergill
c	43	17	28.3	10927	10	AB053447	AB053447 Rattus no
c	44	17	28.3	32557	3	AC006835	AC006835 Caenorhab
c	45	17	28.3	45046	9	AL160000	AL160000 Human DNA

ALIGNMENTS

RESULT 1	BD006937	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006937				
DEFINITION	Artificial promoter libraries for selected organisms and promoters				
ACCESSION	BD006937				
VERSION	BD006937.1	GI:18635308			
KEYWORDS	UP 2001503249-A/6.				
SOURCE	UP 2001503249-A/6.				
ORGANISM	Lactococcus lactis				
REFERENCE	1 (bases 1 to 60)				
AUTHORS	Hammer, K. and Janssen, P. R.				
TITLE	Artificial promoter libraries for selected organisms and promoters				
JOURNAL	Patent: UP 2001503249-A 6 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
OS	Lactococcus lactis				
PN	UP 2001503249-A/6				
PD	13-MAR-2001				
PF	25-AUG-1997 JP 1998510287				
PR	23-AUG-1996 DK 0886/96				
PI	KALIN HAMMER, PETER RUDAR JANSSEN				
PC	CI2N15/09, CI2N15/00				
CC	Strandedness: Double;				
FT	Topology: linear;				
FT	Key				
FT	promoter				
FEATURES	Location/Qualifiers				
source	1..60				
ORIGIN	/organism="Lactococcus lactis"				
ORIGIN	/mol_type="genomic DNA"				
ORIGIN	/db_xref="taxon:1358"				
Query Match	100.0%	Score 60:	DB 6:	Length 60:	
Best Local Similarity	100.0%	Pred. No. 5.1e-23:			
Matches	60:	Conservative	0:	Mismatches	0:
Matches	60:	Indels	0:	Gaps	0:
1	CANGGCTAGTTATCTCTGACAGGAGTATCACTGATATATAGACAGTACTGTT	60			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1662.095 Million cell updates/sec

Title: US-09-242-657b-6

Sequence: 1 CATGGCTTACTTATCTCTG.....TATATAGACAGTACTCTT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_nc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g861.*
9: gb_g862.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	35.0	261	9	CE259388 tigr-g86-
C 2	20	33.3	546	8	B2229259 CH230-488
C 3	20	33.3	703	9	CC914237 t081n22ba
C 4	18	30.0	339	8	AQ085381 HS.21c4.B
C 5	18	30.0	502	8	BH009769 eg25e04.x
C 6	18	30.0	612	6	AI357220 qrx7907.x
C 7	18	30.0	604	6	CA387595
C 8	18	30.0	709	7	CK545518 rswbdo_01
C 9	18	30.0	928	9	CNS06GNY
C 10	18	30.0	1107	8	CC206750 CH261-168
C 11	18	30.0	1221	8	CC233621 CH261-66C
C 12	17	28.3	321	1	AI211642 q0e1a1.f
C 13	17	28.3	442	2	BF615795
C 14	17	28.3	470	8	A2154040 SP.0030.A
C 15	17	28.3	519	1	AV741881
C 16	17	28.3	531	8	AQ783507 HS.3121.A
C 17	17	28.3	570	7	CK501421 rswgao_00
C 18	17	28.3	577	7	CK500601 rswbdo_01
C 19	17	28.3	581	9	CE021053
C 20	17	28.3	585	9	CE021053
C 21	17	28.3	612	9	CE078373 tigr-g86-
C 22	17	28.3	625	9	CK552605 rsw1a0_00
C 23	17	28.3	626	8	AQ418811 RPCI-11-2
C 24	17	28.3	628	7	CF416219

C 25	17	28.3	646	8	AQ510836
C 26	17	28.3	649	2	BE491185
C 27	17	28.3	663	7	CN625364 tae01a11.
C 28	17	28.3	699	8	B2000494 oef14a11.
C 29	17	28.3	709	4	BG645785 EST507404
C 30	17	28.3	715	9	CE079248
C 31	17	28.3	771	7	CF547516
C 32	17	28.3	775	8	AQ738647
C 33	17	28.3	817	9	BX170983
C 34	17	28.3	827	9	BX206808
C 35	17	28.3	838	8	CC322836
C 36	17	28.3	847	9	AZ669235
C 37	17	28.3	854	9	EX137003
C 38	17	28.3	901	8	AZ139626
C 39	17	28.3	934	5	EX851906
C 40	17	28.3	1044	9	CL057021
C 41	17	28.3	1064	8	CC302017
C 42	17	28.3	1106	9	CL057234
C 43	17	28.3	1148	9	CC227553
C 44	17	28.3	1934	8	BH770983
C 45	16	26.7	185	9	AG203415

ALIGNMENTS

RESULT 1
CE259388/c
LOCUS
DEFINITION
tigr-g86-dog-17000346345759 Dog Library Canis familiaris genomic.
ACCESSION
CE259388
VERSION
CE259388.1 GI:35967739
KEYWORDS
GSS.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
REFERENCE
Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K., Ruesch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and Venter, J.C.
AUTHORS
The dog genome: Survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
JOURNAL
MEDLINE
PUBMED
14512627

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shotgun.
Location/Qualifiers
1. 261
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 35.0%; Score 21; DB 9; Length 261;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 TGATATATAGACAGTACTG 58
DB 186 TGATATATAGACAGTACTG 166

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds

(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-6

Perfect score: 60
Sequence: 1 CATGCGCTAGTTATTCTTG.....TATTAATGACAGACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23178 Lactococc
2	18	30.0	58	6	ABK98625 L. lactis
3	18	30.0	58	9	ACD13876 L. lactis
4	18	30.0	64	6	ABK98604 L. lactis
5	18	30.0	64	9	ACD13855 L. lactis
6	18	30.0	65	6	ABK98605 L. lactis
7	18	30.0	65	9	ACD13856 L. lactis
8	18	30.0	93	6	ABK98615 L. lactis
9	18	30.0	93	9	ACD13866 L. lactis
10	18	30.0	9417	6	ABK98632 Vector pr
11	18	30.0	9417	9	ACD13883 L. lactis
12	17	28.3	58	2	AAV23184 Lactococc
13	17	28.3	60	2	AAV23213 Lactococc
14	17	28.3	60	2	AAV23205 Lactococc
15	16	26.7	59	2	AAV23209 Lactococc
16	16	26.7	60	2	AAV23203 Lactococc
17	16	26.7	60	2	AAV23207 Lactococc
18	16	26.7	60	2	AAV23177 Lactococc
19	16	26.7	60	2	AAV23188 Lactococc
20	16	26.7	60	2	AAV23191 Lactococc

21	16	26.7	60	2	AAV23197 Lactococc
22	16	26.7	60	2	AAV23189 Lactococc
23	16	26.7	60	2	AAV23214 Lactococc
24	16	26.7	60	2	AAV23204 Lactococc
25	16	26.7	60	2	AAV23198 Lactococc
26	16	26.7	60	2	AAV23212 Lactococc
27	16	26.7	207	6	ABK98627 L. lactis
28	16	26.7	207	6	ACD13878 L. lactis
29	16	26.7	242	6	ABK98617 L. lactis
30	16	26.7	30	2	ACD13868 L. lactis
31	16	26.7	32	5	ABK15514 Human ner
32	16	26.7	755	6	ABN99221 Arabidops
33	16	26.7	1062	6	ABZ13652 Arabidops
34	16	26.7	1062	6	ADG87654 A. thalia
35	16	26.7	1062	6	ADG87655 A. thalia
36	16	26.7	1062	8	ADG87653 Arabidops
37	16	26.7	1305	3	AAK47953 Arabidops
38	16	26.7	1305	3	AAK51562 Arabidops
39	16	26.7	1684	2	AAZ27540 Wild type
40	16	26.7	1689	2	AAZ27539 Truncated
41	16	26.7	2000	12	ADJ41509 Plant cDN
42	16	26.7	2118	3	AAK51464 Arabidops
43	16	26.7	3980	2	AAQ36780 Human car
44	16	26.7	3980	13	ADR48957 Human pho
45	16	26.7	4306	8	ABZ81660 Human typ

ALIGNMENTS

RESULT 1	AAV23178 standard; DNA; 60 BP.
ID	AAV23178; (first entry)
AC	28-JUL-1998
XX	Lactococcus lactis constitutional promoter Cp10.
DE	Lactococcus lactis; constitutional promoter; optimise; spacer;
KW	artificial promoter library; gene expression; ds.
XX	Synthetic.
OS	Lactococcus lactis.
XX	Key
XX	Location/Qualifiers
XX	4..60
XX	/tag= a
XX	/standard_name= "Constitutional promoter"
XX	WO9807846-A1.
XX	26-FEB-1998.
XX	25-AUG-1997; 97WO-DK000342.
XX	23-AUG-1996; 96DK-00000886.
XX	(JENS/) JENSEN P R.
XX	Jensen PR, Hammer K;
XX	WPI, 1998-179062/16;
XX	New artificial promoter libraries - containing consensus promoter
XX	sequences and variable spacers, used to generate promoters for optimising
XX	expression of genes.
XX	Claim 28; Page 43; 89pp; English.
XX	This is a Lactococcus lactis constitutional promoter sequence used in the
XX	construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Genomic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 / Search time 131.858 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-7

Perfect score: 59
Sequence: 1 CATAGTCAGTTATCTCTG.....ATATTAATAGTACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	2	AAV23179
2	17	28.8	60	2	AAV23189
3	17	28.8	3136	3	AAV16631
4	17	28.8	3140	5	ABA82682
5	17	28.8	3140	8	ACC46002
6	17	28.8	3140	10	ADB89696
7	17	28.8	3140	10	ADBE2492
8	17	28.8	3227	4	AAF59596
9	17	28.8	12781	10	ADBS3970
10	17	28.8	13275	8	ABZ74033
11	17	28.8	13275	8	ADA98640
12	17	28.8	13275	10	ADC20763
13	17	28.8	13275	10	ABZ67620
14	17	28.8	58857	3	AAA58471
15	17	27.1	59	2	AAV23185
16	17	27.1	265	2	AAQ45205
17	17	27.1	265	2	AAQ45205
18	17	27.1	265	2	AAQ45205
19	16	27.1	314	6	ABK62419
20	16	27.1	314	10	ADBS5788

21	16	27.1	314	10	ABT40651
22	16	27.1	314	12	ADP71740
23	16	27.1	418	9	ACH17078
24	16	27.1	645	9	AAI19008
25	16	27.1	1011	8	ACA27756
26	16	27.1	1219	4	AAH72911
27	16	27.1	1258	6	ABK36209
28	16	27.1	1284	3	AAQ99017
29	16	27.1	1782	6	ABBS2589
30	16	27.1	1851	6	ABK3968
31	16	27.1	1851	13	ADR99063
32	16	27.1	2220	4	AAQ26157
33	16	27.1	2864	12	ADP50636
34	16	27.1	3814	11	ADP44917
35	16	27.1	3945	10	ADP81655
36	16	27.1	12781	4	AA546386
37	16	27.1	12781	6	ABU92229
38	16	27.1	12781	10	ADBS4256
39	16	27.1	12781	10	ADBS4128
40	16	27.1	15382	8	ACD13393
41	16	27.1	24389	4	ABU29006
42	16	27.1	28046	6	ABK69837
43	16	27.1	43591	11	ACN44916
44	16	27.1	97658	8	ABQ83210_3
45	16	27.1	110000	6	ABA90521_02

ALIGNMENTS

RESULT 1	AAV23179	standard; DNA; 59 BP.
ID	AAV23179;	
AC	AAV23179;	
XX		
DT	28-JUL-1998	(first entry)
XX		
DE	Lactococcus lactis constitutional promoter Cpl1.	
XX		
KW	Lactococcus lactis; constitutional promoter; optimum; spacer;	
XX	artificial promoter library; gene expression; ds.	
OS	Synthetic.	
XX	Lactococcus lactis.	
XX		
PH	Key	Location/Qualifiers
FT	Promoter	4..59
FT		/*tag= a
XX		/standard_name= "Constitutional promoter"
XX		
XX	W09807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DX-00000886.
XX		
PA	(JENSEN) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
PS	Claim 28; Page 44; 89pp; English.	
XX		
CC	This is a Lactococcus lactis constitutional promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 458.474 Seconds
(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-7

Perfect score: 59
Sequence: 1 CATAGTACGATTATCTCTG.....ATATATAGTACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenBdb1.*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_lm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_um:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	59	100.0	59	BD006938	Artificia
2	20	33.9	90175	AC092391	AC092391 Homo sapi
3	20	33.9	111799	HS1850015	HS1850015 Human DNA
4	20	33.9	174874	AC026259	AC026259 Homo sapi
5	19	32.2	230236	AC093366	AC093366 Mus muscu
6	18	30.5	89159	AL593845	AL593845 Zebrafish
7	18	30.5	105161	AC144728	AC144728 Medicago
8	18	30.5	106528	AC136840	AC136840 Medicago
9	18	30.5	110000	1 AE017197	1 AE017197 03
10	18	30.5	110000	5 AC111021	5 AC111021 0
11	18	30.5	153540	5 BX324186	5 BX324186 Mus muscu
12	18	30.5	153539	5 AL953846	5 AL953846 Zebrafish
13	18	30.5	173414	9 AC010387	9 AC010387 Homo sapi
14	18	30.5	185531	9 AC012059	9 AC012059 Homo sapi
15	18	30.5	188892	2 AC091327	2 AC091327 Mus muscu
16	18	30.5	192652	2 BX927230	2 BX927230 Danio rer
17	18	30.5	215810	10 AC120788	10 AC120788 Mus muscu
18	18	30.5	222345	2 AC137233	2 AC137233 Rattus no
19	18	30.5	248046	2 AC122667	2 AC122667 Rattus no

c	20	18	30.5	250156	2	AC094168	AC094168 Rattus no
c	21	17	28.8	60	6	BD006948	BD006948 Artificia
c	22	17	28.8	173	8	S4620982	S4620982 triose phos
c	23	17	28.8	400	6	CQ709958	CQ709958 Sequence
c	24	17	28.8	667	11	BV030702	BV030702 S212P6016
c	25	17	28.8	693	5	AY223644	AY223644 Lachesis mu
c	26	17	28.8	694	5	LMU41885	LMU41885 Lachesis mu
c	27	17	28.8	3122	6	CQ728783	CQ728783 Sequence
c	28	17	28.8	3128	9	BC017713	BC017713 Homo sapi
c	29	17	28.8	3136	6	BD227253	BD227253 Secreted
c	30	17	28.8	3140	6	AX277589	AX277589 Sequence
c	31	17	28.8	3140	9	AF053977	AF053977 Homo sapi
c	32	17	28.8	3227	6	AX078257	AX078257 Sequence
c	33	17	28.8	3848	1	STRSTRA	M26130 S. parasang
c	34	17	28.8	12781	6	AX822134	AX822134 Sequence
c	35	17	28.8	12781	6	AX825774	AX825774 Sequence
c	36	17	28.8	20125	1	AF461770	AF461770 Yersinia
c	37	17	28.8	24979	5	AL672198	AL672198 Zebrafish
c	38	17	28.8	26757	9	AY082894	AY082894 Homo sapi
c	39	17	28.8	28870	9	AY603103	AY603103 Homo sapi
c	40	17	28.8	34758	9	AL357147	AL357147 Human DNA
c	41	17	28.8	71075	9	AC109442	AC109442 Homo sapi
c	42	17	28.8	77457	1	AF210249	AF210249 Streptomy
c	43	17	28.8	79046	8	AB008267	AB008267 Arabidops
c	44	17	28.8	79730	2	AC023036	AC023036 Homo sapi
c	45	17	28.8	87027	2	BX539319	BX539319 Danio rer

ALIGNMENTS

RESULT 1	BD006938	59 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Artificial promoter libraries for selected organisms and promoters derived from such libraries.				
ACCESSION	BD006938	1	GI:18635309		
VERSION	JP 2001503249-A/7.				
KEYWORDS	Lactococcus lactis				
SOURCE	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
ORGANISM	1 (bases 1 to 59)				
REFERENCE	Hammer, K. and Janssen, P.R.				
AUTHORS	Artificial promoter libraries for selected organisms and promoters derived from such libraries				
TITLE	Patent: JP 2001503249-A 7 13-MAR-2001;				
JOURNAL	PETER RUDAR JANSSEN				
COMMENT	OS Lactococcus lactis				
	PN JP 2001503249-A/7				
	PD 13-MAR-2001				
	PE 25-AUG-1997 JP 1998510287				
	PR 23-AUG-1996 DK 0886/96				
	PI KALIN HAMMER, PETER RUDAR JANSSEN				
	PC C12N15/09, C12N15/00				
	CC Strandedness: Double;				
	CC Topology: Linear;				
	CH Key				
	FT promoter				
FEATURES	Location/Qualifiers				
source	1..59				
	/organism="Lactococcus lactis"				
	/mol_type="Genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN	Query Match 100.0%; Score 59; DB 6; Length 59;				
	Best Local Similarity 100.0%; Pred. No. 1.1e-23;				
	Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 CATAGTACGATTATCTTACCCGACGCCCCCTTGATATATAGTACTGTT 59				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1206.06 Seconds

(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-7

Perfect score: 59
Sequence: 1 CATAGAGATTATTCTTC.....ATATATAGTACTACTGTT 59

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	32.2	642	2	BB483350 BB483350
2	19	32.2	670	2	BB650150 BB650150
3	19	32.2	781	2	CR231312 Forward s
4	19	32.2	899	9	CR024658 Forward s
5	19	32.2	911	9	CR104784 Forward s
6	19	32.2	1289	5	BM910149 AGENCOURT
7	19	32.2	2811	3	AK084963 Mus muscu
8	19	32.2	3523	3	AK082491 Mus muscu
9	19	32.2	329	6	CD584638 RK027A1FO
10	18	30.5	330	6	CD596688 RK103A3CO
11	18	30.5	521	6	CD052758 CD052758
12	18	30.5	539	6	CB360664 ZF001-P00
13	18	30.5	575	6	CB352534 ZF001-P00
14	18	30.5	646	6	AI1942750 fct7c01.Y
15	18	30.5	650	6	CD593903 RK088A3HO
16	18	30.5	650	6	CD595455 RK097A1AO
17	18	30.5	687	2	CF544042 DREXAM-56
18	18	30.5	687	2	AM165307 CD052694
19	18	30.5	775	6	CD052694 LITHZP000
20	18	30.5	955	7	CO007659 EST795994
21	18	30.5	980	7	CO007658 EST795993
22	18	30.5	1016	9	CL017663 CH216-2H1
23	18	30.5	1101	9	AL007158 Drosophila
24	18	30.5	1115	5	BB677269 AGRNCOURT

25	17	28.8	168	2	BE765119 CM3-NT009
26	17	28.8	294	2	BE091438 PM3-BT072
27	17	28.8	296	2	BE091401 PM4-BT072
28	17	28.8	326	1	AA312999 EST183790
29	17	28.8	349	2	AA022483 df39f12.Y
30	17	28.8	354	8	B2287254 SALK_0206
31	17	28.8	367	2	AM891861 CM3-NT009
32	17	28.8	370	2	AM959684 EST371754
33	17	28.8	372	2	BF477745 BPO81174
34	17	28.8	376	5	BP081174 BPO81174
35	17	28.8	395	2	AM891863 CM3-NT009
36	17	28.8	417	2	AM891676 CM3-NT009
37	17	28.8	418	2	AM891858 CM3-NT009
38	17	28.8	421	6	CB798179 AMGNNUC:S
39	17	28.8	422	2	AM891865 CM3-NT009
40	17	28.8	430	2	AM891685 CM3-NT009
41	17	28.8	443	1	AA165669 Z080A05.T
42	17	28.8	445	1	AA159602 Z080A05.T
43	17	28.8	446	1	AJ653066 AJ653066
44	17	28.8	461	2	BF887236 CM4-TN014
45	17	28.8	483	8	AQ726829 HS_5332_B

ALIGNMENTS

RESULT 1
LOCUS BB483350 642 bp mRNA linear EST 25-OCT-2001
DEFINITION BB483350 RIKEN full-length enriched, 13 days embryo lung Mus
ACCESSION BB483350
VERSION BB483350.2 GI:16441028
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

Akakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Komno, H., Kouda, M., Koya, S., Matsumura, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Akakawa, T., et al. 2001) Unpublished (2001)

On Jul 23, 2000 this sequence version replaced gi:9400959.
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsutani, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsutani, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Komno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-8

Perfect score: 60
Sequence: 1 CATATCAACTTATCTCTG.....TATATACCTGAGTACTGTT 60

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	AAV23180	Aav23180 Lactococc
2	19	31.7	496	AAI83884	Aai83884 Human pol
3	19	31.7	651	AAFI3601	Aafi3601 Aspergill
4	19	31.7	1543	ADRE1001	Adre1001 Cotton CD
5	19	31.7	3495	ABLO2168	Ablo2168 Drosophil
6	18	30.0	316	ABN95301	Abn95301 Gene #179
7	18	30.0	343	AAI07992	Aai07992 Human bre
8	18	30.0	388	AAI25738	Aai25738 Human bre
9	18	30.0	440	AAI18133	Aai18133 Human bre
10	18	30.0	610	AAI26781	Aai26781 Human bre
11	18	30.0	1508	AAI45947	Aai45947 Human bre
12	18	30.0	1779	ACN88719	Acn88719 Breast ca
13	18	30.0	2774	ADM02497	Adm02497 Human CDN
14	18	30.0	3344	AAH15978	Aah15978 Human CDN
15	18	30.0	3484	AAAC1047	Aaac1047 Human sec
16	18	30.0	4836	ACH04259	Acho4259 Human sec
17	18	30.0	5662	ACN92609	Acn92609 Breast ca
18	18	30.0	162450	AAZ86967	Aaz86967 Retinobla
19	17	28.3	60	AAV23207	Aav23207 Lactococc
20	17	28.3	60	AAV23208	Aav23208 Lactococc

21	17	28.3	2864	12	ADF50636	Adf50636 Murine DN
22	16	26.7	59	2	AAV23190	Aav23190 Lactococc
23	16	26.7	60	2	AAV23211	Aav23211 Lactococc
24	16	26.7	60	2	AAV23203	Aav23203 Lactococc
25	16	26.7	60	2	AAV23182	Aav23182 Lactococc
26	16	26.7	60	2	AAV23197	Aav23197 Lactococc
27	16	26.7	60	2	AAV23214	Aav23214 Lactococc
28	16	26.7	60	2	AAV23204	Aav23204 Lactococc
29	16	26.7	60	2	AAV23205	Aav23205 Lactococc
30	16	26.7	496	6	ABR44927	Abbr44927 CDNA enco
31	16	26.7	557	6	ABN64949	Abn64949 Human can
32	16	26.7	621	5	ABV52236	Abv52236 Human pro
33	16	26.7	676	2	AAI26920	Aai26920 Human gen
34	16	26.7	797	2	AAV36459	Aav36459 caryabl DN
35	16	26.7	1056	6	AAI46978	Aai46978 Cell cycl
36	16	26.7	1233	6	AAI46971	Aai46971 Human cel
37	16	26.7	1275	6	AAI46969	Aai46969 Human cel
38	16	26.7	1329	6	AAI46970	Aai46970 Human cel
39	16	26.7	1350	5	ABN02875	Abn02875 Mouse SKP
40	16	26.7	1410	9	ADA28837	Ada28837 DNA encod
41	16	26.7	1462	9	ADA14305	Ada14305 Human SKP
42	16	26.7	1462	13	ACN38212	Acn38212 Tumour-as
43	16	26.7	1561	4	ACB84601	Acb84601 Beta-TRCP
44	16	26.7	1600	2	AAI62159	Aai62159 S-phase K
45	16	26.7	1600	4	AAI62159	Aai62159 S-phase K

ALIGNMENTS

RESULT 1	AAV23180	standard; DNA; 60 BP.
ID	AAV23180	standard; DNA; 60 BP.
XX	AAV23180;	
AC	XX	
DT	28-JUL-1998	(first entry)
XX	XX	
DE	Lactococcus lactis	constitutional promoter Cpl2.
XX	XX	
KW	Lactococcus lactis;	constitutional promoter; optimise; spacer;
XX	artificial promoter library;	gene expression; ds.
OS	Synthetic.	
OS	Lactococcus lactis.	
XX	XX	
FT	Key	Location/Qualifiers
FT	promoter	4..60
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
PN	W09807846-A1.	
XX	XX	
PD	26-FEB-1998.	
XX	XX	
PF	25-AUG-1997;	97WO-DK000342.
XX	XX	
PR	23-AUG-1996;	96DK-00000886.
XX	XX	
PA	(JENSEN) JENSEN P R.	
XX	XX	
FI	Jensen PR, Hammer K;	
XX	XX	
DR	WPI, 1998-179062/16.	
XX	XX	
PT	New artificial promoter libraries -	containing consensus promoter
PT	sequences and variable spacers,	used to generate promoters for optimising
XX	expression of genes.	
PS	Claim 28; Page 44; 89pp; English.	
XX	XX	
CC	This is a Lactococcus lactis	constitutional promoter sequence used in the
CC	construction of an artificial	promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657b-8

Perfect score: 60

Sequence: 1 CATATACAGTTATCTCTG.....TATATACCTGACTGTT 60

Scoring table: OLIGO_NUC

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenBank: 1: gb ba: *
2: gb prt: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb pat: *
7: gb ph: *
8: gb pl: *
9: gb pr: *
10: gb_ro: *
11: gb_sbs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006939	Artificial promoter libraries for selected organisms and promoters
2	20	33.3	177092	AC119446	Artificial promoter libraries for selected organisms and promoters
3	20	33.3	245468	AC130746	Artificial promoter libraries for selected organisms and promoters
4	19	31.7	3495	CQ572735	Artificial promoter libraries for selected organisms and promoters
5	19	31.7	38193	AC017671	Artificial promoter libraries for selected organisms and promoters
6	19	31.7	139121	BX511186	Artificial promoter libraries for selected organisms and promoters
7	19	31.7	140266	AC124406	Artificial promoter libraries for selected organisms and promoters
8	19	31.7	144511	BX927332	Artificial promoter libraries for selected organisms and promoters
9	19	31.7	158504	AC107810	Artificial promoter libraries for selected organisms and promoters
10	19	31.7	167977	AC010010	Artificial promoter libraries for selected organisms and promoters
11	19	31.7	168250	AC093121	Artificial promoter libraries for selected organisms and promoters
12	19	31.7	197721	AL954771	Artificial promoter libraries for selected organisms and promoters
13	19	31.7	212445	AC108419	Artificial promoter libraries for selected organisms and promoters
14	19	31.7	300610	AE003467	Artificial promoter libraries for selected organisms and promoters
15	19	31.7	323782	BX510056	Artificial promoter libraries for selected organisms and promoters
16	19	30.0	316	AX409152	Artificial promoter libraries for selected organisms and promoters
17	19	30.0	343	CQ415415	Artificial promoter libraries for selected organisms and promoters
18	19	30.0	388	CQ433174	Artificial promoter libraries for selected organisms and promoters
19	19	30.0	440	CQ425566	Artificial promoter libraries for selected organisms and promoters

20	18	30.0	610	6	CQ434218	Sequence
21	18	30.0	2770	6	BD135220	Human nuc
22	18	30.0	2770	6	AX017601	Sequence
23	18	30.0	2774	6	AX834058	Sequence
24	18	30.0	2774	6	AK096270	Sequence
25	18	30.0	3344	6	BD157970	Primer fo
26	18	30.0	3344	6	AX879689	Sequence
27	18	30.0	3344	6	AK022729	Sequence
28	18	30.0	3548	9	AB053320	Homo sapi
29	18	30.0	3548	9	AB053320	Homo sapi
30	18	30.0	48126	9	AJ584659	Homo sapi
31	18	30.0	55198	9	AL136083	Homo sapi
32	18	30.0	58072	2	AC136758	Human DNA
33	18	30.0	65716	10	AC003657	Homo sapi
34	18	30.0	106482	10	AC007585	Homo sapi
35	18	30.0	124148	10	AC134433	Mus muscu
36	18	30.0	142273	9	AL139194	Human DNA
37	18	30.0	154935	2	AL136781	Homo sapi
38	18	30.0	155152	2	AL139808	Homo sapi
39	18	30.0	160288	10	AC006584	Homo sapi
40	18	30.0	162238	10	AC145392	Mus muscu
41	18	30.0	162450	6	BD221862	Nucleic a
42	18	30.0	162450	6	AR211792	Sequence
43	18	30.0	171898	2	CR388126	Danio rer
44	18	30.0	173048	5	AL772154	Zebrafish
45	18	30.0	174074	9	AC018891	Homo sapi

ALIGNMENTS

RESULT 1	BD006939	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006939	Artificial promoter libraries for selected organisms and promoters			
DEFINITION	BD006939	Artificial promoter libraries for selected organisms and promoters			
ACCESSION	BD006939.1	GI:18635310			
VERSION	JP 2001503249-A/8.				
KEYWORDS	Lactococcus lactis				
SOURCE	Lactococcus lactis				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
REFERENCE	1 (bases 1 to 60)				
AUTHORS	Hammer, K. and Janssen, P.R.				
TITLE	Artificial promoter libraries for selected organisms and promoters				
JOURNAL	Patent: JP 2001503249-A 8 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
OS	Lactococcus lactis				
PN	JP 2001503249-A/8				
PD	13-MAR-2001				
PF	25-AUG-1997 JP 1998510287				
PI	25-AUG-1996 DK 0886/96				
CC	KALIN HAMMER, PETER RUDAR JANSSEN				
CC	CI2N15/09, CI2N15/00				
CC	Strandedness: Double;				
CC	Topology: linear;				
FT	Key				
FT	promoter				
FEATURES	Location/Qualifiers				
source	1..60				
	/organism="Lactococcus lactis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN					
Query Match	100.0%; Score 60; DB 6; Length 60;				
Best Local Similarity	100.0%; Pred. No. 3.8e-23;				
Matches	60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 CATATACAGTTATCTGACACTGACGCGCAAAATGATATTAACCTGACTGTT 60				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-8

Perfect score: 60
Sequence: 1 CATATACAGTTCCTTCTTG.....TATTAATCCTGAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapect 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	443	8	AQ458484 HS_5061_A
2	19	31.7	644	7	CO151372 EST826425
3	19	31.7	673	7	CO152362 EST827415
4	18	30.0	287	1	AA327610 EST30992
5	18	30.0	316	7	F09058
6	18	30.0	332	5	BQ364238
7	18	30.0	371	1	AA702576
8	18	30.0	396	5	BQ708953
9	18	30.0	405	7	H49083
10	18	30.0	409	1	AA421057
11	18	30.0	422	1	AA1368772
12	18	30.0	425	2	BF445392
13	18	30.0	434	1	AA889598
14	18	30.0	446	2	BF114947
15	18	30.0	455	1	AI073358
16	18	30.0	465	6	CB163652
17	18	30.0	514	1	AV738448
18	18	30.0	514	7	H10450
19	18	30.0	538	8	CR079918
20	18	30.0	561	9	CR079918
21	18	30.0	641	1	AV658411
22	18	30.0	648	2	AW959911
23	18	30.0	695	5	BU673992
24	18	30.0	729	2	BF674940

25	18	30.0	772	7	CN501862	CN501862
26	18	30.0	823	8	BZ785137	BZ785137
27	18	30.0	870	7	CO933162	CO933162
28	18	30.0	874	9	EX985571	EX985571
29	18	30.0	877	8	BZ045644	BZ045644
30	17	28.3	217	9	CR046864	CR046864
31	17	28.3	236	2	BB147650	BB147650
32	17	28.3	285	2	AW255399	AW255399
33	17	28.3	377	8	AO106268	AO106268
34	17	28.3	450	6	CD096291	CD096291
35	17	28.3	452	9	CE060197	CE060197
36	17	28.3	477	8	AQ774538	AQ774538
37	17	28.3	512	8	AQ213073	AQ213073
38	17	28.3	530	8	AQ438452	AQ438452
39	17	28.3	553	2	AW202959	AW202959
40	17	28.3	565	4	BM154642	BM154642
41	17	28.3	577	7	CK088146	CK088146
42	17	28.3	603	6	CD096238	CD096238
43	17	28.3	622	7	CF236236	CF236236
44	17	28.3	627	6	CA070930	CA070930
45	17	28.3	628	9	CE335717	CE335717

ALIGNMENTS

RESULT 1
LOCUS: AQ458484
DEFINITION: HS_5061_A1_C04_SPEB_RPCI-11 Human Male BAC library Homo sapiens genomic clone Plate=637 Col=7 Row=5, genomic survey sequence.
ACCESSION: AQ458484
VERSION: AQ458484.1 GI:4637124
KEYWORDS: GSS.
SOURCE: Homo sapiens (human)
ORGANISM: Homo sapiens
REFERENCE: 1 (bases 1 to 443)
AUTHORS: Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Purlong J., Young J., Zhao S., Adams W.D. and Hood L.
TITLE: Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL: Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE: 99380589
PUBMED: 10449764

COMMENT: Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Peter de Jong (peterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 637 row: 5 column: 7
Seg primer: SP6
Class: BAC ends
High quality sequence stop: 443.

FEATURES

source

Location/Qualifiers
1..443
/Organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=637 Col=7 Row=5"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657b-9

Perfect score: 60
Sequence: 1 CAGGCTTACTTATCTCTG.....TATATACGTGAACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1808:*
2: geneseqn1908:*
3: geneseqn2008:*
4: geneseqn2018:*
5: geneseqn2018b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2 AAV23181	AAV23181 Lactococc
2	18	30.0	22259	4 AAL04376	AAL04376 Human rep
3	18	30.0	27804	4 AAK86476	AAK86476 Human imm
4	17	28.3	417	6 ABN73940	ABN73940 Bovine em
5	17	28.3	821	6 AAS62162	AAS62162 Porcine m
6	17	28.3	4428	4 AAD06578	AAD06578 Porcine a
7	17	28.3	4428	4 AAD06574	AAD06574 Bovine al
8	17	28.3	4428	4 AAD06575	AAD06575 Bovine al
9	17	28.3	63000	6 ABS67634	ABS67634 Human cab
10	17	28.3	78056	8 ABV99701	ABV99701 Bovine BS
11	16	26.7	115	6 ABK98626	ABK98626 L. lactis
12	16	26.7	115	9 ACD13877	ACD13877 L. lactis
13	16	26.7	150	6 ABK98616	ABK98616 L. lactis
14	15	26.7	150	9 ACD13867	ACD13867 L. lactis
15	15	26.7	170	12 ACH92995	ACH92995 Human gen
16	15	26.7	180	4 ABA73456	ABA73456 Human foe
17	15	26.7	180	4 AAI53890	AAI53890 Probe #22
18	16	26.7	180	4 ABA47776	ABA47776 Human liv
19	16	26.7	207	6 ABK98627	ABK98627 L. lactis
20	16	26.7	207	9 ACD13878	ACD13878 L. lactis

21	16	26.7	242	6 ABK98617	ABK98617 L. lactis
22	16	26.7	242	9 ACD13868	ACD13868 L. lactis
23	16	26.7	462	5 AAS74788	AAS74788 DNA encod
24	16	26.7	500	12 ACH79295	ACH79295 Human gen
25	16	26.7	572	4 AAI17339	AAI17339 Probe #72
26	16	26.7	572	4 ABA62255	ABA62255 Human foe
27	16	26.7	572	4 AAI42233	AAI42233 Probe #10
28	16	26.7	572	4 ABA29601	ABA29601 Probe #80
29	16	26.7	572	4 AAK36483	AAK36483 Human don
30	16	26.7	572	4 AAK10588	AAK10588 Human bra
31	16	26.7	572	4 ABA36133	ABA36133 Human liv
32	16	26.7	572	6 ABA10480	ABA10480 Human gen
33	16	26.7	587	4 ABA60930	ABA60930 Human foe
34	16	26.7	587	4 AAI40825	AAI40825 Probe #95
35	16	26.7	587	4 ABA34861	ABA34861 Human liv
36	16	26.7	634	5 AAS80637	AAS80637 DNA encod
37	16	26.7	638	5 AAS74790	AAS74790 DNA encod
38	16	26.7	842	12 ADJ39761	ADJ39761 Plant cDN
39	16	26.7	1131	10 ADG48196	ADG48196 Mycobacte
40	16	26.7	1155	6 ABN92369	ABN92369 Stephyloc
41	16	26.7	1155	13 ADS03068	ADS03068 Stephyloc
42	16	26.7	1156	6 ABR74359	ABR74359 Bacillus
43	16	26.7	1553	4 AAS41585	AAS41585 cDNA enco
44	16	26.7	1642	13 ADR24458	ADR24458 Breast ca
45	16	26.7	1842	2 AAV84560	AAV84560 Human sec

ALIGNMENTS

RESULT 1	AAV23181	AAV23181 standard; DNA; 60 BP.
XX	AAV23181;	
DT	28-JUL-1998	(first entry)
XX		
DE	Lactococcus lactis	constitutional promoter Cpl3.
XX		
KW	Lactococcus lactis;	constitutional promoter; optimise; spacer;
XX	artificial promoter library;	gene expression; de.
OS	Synthetic.	
XX	Lactococcus lactis.	
XX		
Key		Location/Qualifiers
FT	Promoter	4..60
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
XX		
PN	W09807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENS/) JENSEN P R.	
XX		
PT	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
PS	Claim 28; Page 45; 89pp; English.	
XX		
CC	This is a Lactococcus lactis constitutional promoter sequence used in the	
	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(Without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-9

Perfect score: 60

Sequence: 1 CATGCTTACTTATCTCTG.....TATATACGAGAACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_mt:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006940	Artificial
2	19	31.7	367	D63661	Volvox sp.
3	19	31.7	564	AF308999	Alpheus a
4	19	31.7	124721	AP005575	Oryza sat
5	19	31.7	128635	AC108074	Homo sapi
6	19	31.7	154561	AL929071	Mouse DNA
7	19	31.7	158798	AP005308	Oryza sat
8	19	31.7	172281	AC096531	Homo sapi
9	19	31.7	181208	AP003063	Homo sapi
10	19	31.7	199016	AC021387	Homo sapi
11	19	31.7	200276	AP003100	Homo sapi
12	19	30.0	420	AY028590	Stenase1
13	19	30.0	420	AY028591	Stenase1
14	18	30.0	541	AY344701	Synalphen
15	18	30.0	541	AY344703	Synalphen
16	18	30.0	551	SCJ225879	Sesarma c
17	18	30.0	551	SDJ225879	Sesarma c
18	18	30.0	551	SRJ225877	Sesarma r
19	18	30.0	551	SSJ225863	Sesarma s

20	18	30.0	551	3	SSJ225889	AJ225889 Sesarma s
21	18	30.0	551	3	SVJ225858	AJ225858 Sesarma v
22	18	30.0	551	3	SWJ225886	AJ225886 Sesarma w
23	18	30.0	561	3	AF321335	AF321335 Alpheus a
24	18	30.0	564	3	AF309872	AF309872 Alpheus a
25	18	30.0	564	3	AF321344	AF321344 Alpheus f
26	18	30.0	574	3	AF501661	AF501661 Alpheus a
27	18	30.0	577	3	AY135191	AY135191 Alpheus a
28	18	30.0	592	3	AF230796	AF230796 Synalphen
29	18	30.0	609	3	S65261	S65261 Litopenaeus
30	18	30.0	620	3	AF501656	AF501656 Alpheus a
31	18	30.0	625	3	AF501660	AF501660 Alpheus a
32	18	30.0	625	3	AF501651	AF501651 Alpheus a
33	18	30.0	628	3	AF501659	AF501659 Alpheus a
34	18	30.0	629	3	AF501658	AF501658 Alpheus a
35	18	30.0	630	3	AF501650	AF501650 Alpheus a
36	18	30.0	630	3	AF501657	AF501657 Alpheus a
37	18	30.0	631	3	AF501655	AF501655 Alpheus a
38	18	30.0	640	3	AF501652	AF501652 Alpheus a
39	18	30.0	640	3	AF501653	AF501653 Alpheus a
40	18	30.0	640	3	AF501654	AF501654 Alpheus a
41	18	30.0	847	3	AY264894	AY264894 Parapenae
42	18	30.0	847	3	AY264903	AY264903 Trachypen
43	18	30.0	1430	3	DDCORONIN	X61480 D. discoideu
44	18	30.0	1618	9	HIMPSRG02	M28235 Homo sapien
45	18	30.0	1626	8	AK060989	AK060989 Oryza sat

ALIGNMENTS

RESULT 1
BD006940
LOCUS
DEFINITION Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION BD006940.1 GI:18635311
VERSION JP 2001503249-A/9.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE 1 (bases 1 to 60)
AUTHORS Hammer, K. and Janssen, P. R.
TITLE Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL Patent: JP 2001503249-A 9 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Lactococcus lactis
PN JP 2001503249-A/9
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (4)..(60).
Location/Qualifiers
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 7.8e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGCTTACTTATCTTGAACAAACAGCTTTGTATATACGTGAGAACTGTT 60
|||||

FEATURES PROJECT = Dictyostelium discoideum cDNA project in Japan.

source Location/Qualifiers
1.323

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SLJ246"
/dev_stage="slug"
/clone_lib="Dictyostelium discoideum SL (H.Urushihara) "

ORIGIN

Query Match 30.0%; Score 18; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 37;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TTGACAAACCAACGAGCT 35
DB 131 TTGACAAACCAACGAGCT 114

RESULT 15

AU040066/c

LOCUS AU040066 325 bp mRNA linear EST 29-MAR-1999

DEFINITION AU040066 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium

ACCESSION AU040066

VERSION AU040066.1 GI:4009306

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.

1 (bases 1 to 325)

Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,

Williams, U., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.

The Dictyostelium developmental cDNA project: generation and

analysis of expressed sequence tags from the first-finger stage of

development

DNA Res. 5 (6), 335-340 (1998)

JOURNAL 99156227

MEDLINE 10048482

COMMENT

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp

PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.

Location/Qualifiers

1.325

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="SLA350"

/dev_stage="slug"

/clone_lib="Dictyostelium discoideum SL (H.Urushihara) "

ORIGIN

Query Match 30.0%; Score 18; DB 1; Length 325;

Best Local Similarity 100.0%; Pred. No. 37;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 TTGACAAACCAACGAGCT 35
DB 130 TTGACAAACCAACGAGCT 113

Search completed: June 21, 2005, 05:51:30
Job time : 1230.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds

(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-10

Perfect score: 60
Sequence: 1 CATGACGAGATTATCTCTG.....TATATAAAGACGACGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23182
2	20	33.3	60	2	AAV23211
3	19	31.7	4136	4	AAH54672
4	18	30.0	60	2	AAV23177
5	17	28.3	59	2	AAV23190
6	17	28.3	59	2	AAV23210
7	17	28.3	60	2	AAV23191
8	17	28.3	60	2	AAV23197
9	17	28.3	60	2	AAV23200
10	17	28.3	60	2	AAV23199
11	17	28.3	765	10	ADP02876
12	17	28.3	2864	12	ADP50636
13	16	26.7	59	2	AAV23185
14	16	26.7	60	2	AAV23208
15	16	26.7	60	2	AAV23189
16	16	26.7	60	2	AAV23214
17	16	26.7	60	2	AAV23198
18	16	26.7	60	2	AAV23180
19	16	26.7	60	2	AAV23205
20	16	26.7	265	2	AAQ23880

21	16	26.7	265	2	AAQ46205	AaQ46205 Promoter
22	16	26.7	265	2	AAT31875	Aat31875 Promoter
23	16	26.7	296	3	AAC03364	Aac03364 Human sec
24	16	26.7	363	10	ADD19955	Add19955 Oreochrom
25	16	26.7	651	3	AAI13601	Aai13601 Aspergill
26	16	26.7	1266	13	ADSS8205	Adss8205 Bacterial
27	16	26.7	1779	12	ADP81023	Adp81023 Human ova
28	16	26.7	1801	6	ABZ57614	Abz57614 Human aut
29	16	2007	12	ADO57332	Ado57332 DNA encod	
30	16	28.7	2892	4	AAH16551	Aah16551 Human cDN
31	16	26.7	12537	2	AAT41705	Aat41705 Lymphocyt
32	16	26.7	13635	4	ABU11970	Abu11970 Drosophil
33	16	26.7	15755	12	ADM66947	Adm66947 Murine ad
34	16	26.7	37672	8	AAD56120	Aad56120 Mouse irf
35	16	26.7	37672	9	ADA02483	Ada02483 Mouse irf
36	16	26.7	37672	10	ADB72221	Adb72221 Mouse irf
37	16	26.7	38239	12	ADQ97626	Adq97626 Mouse can
38	16	26.7	54701	11	ACN44478	Acn44478 Human gen
39	16	26.7	81463	12	ADQ97682	Adq97682 Mouse can
40	16	26.7	86248	10	ADC00087	Adc00087 Enterohae
41	16	26.7	87563	9	ACD19044	Accl19044 E. coli 0
42	16	26.7	98865	6	ABQ78054	Abq78054 Human Ras
43	16	26.7	104399	13	ABD33148	Abd33148 Murine ca
44	16	26.7	110000	6	ABA90521_02	Continuation (3 of
45	16	26.7	110000	6	ABA90521_03	Continuation (4 of

ALIGNMENTS

RESULT 1
ID AAV23182 standard; DNA; 60 BP.
XX AAV23182;

DT 28-UTL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cpl4.

KW Lactococcus lactis; constitutional promoter; optimise; spacer;

OS Synthetic.

OS Lactococcus lactis.

PH Key Location/Qualifiers

FT promoter 4..60 /tag= a

FT promoter /standard_name= "Constitutional promoter"

PN WO9807846-A1.

PD 26-FEB-1998.

PF 25-AUG-1997; 97WO-DK000342.

PR 23-AUG-1996; 96DK-00000886.

PA (JENSEN) JENSEN P R.

PI Jensen PR, Hammer K;

DR WPI, 1998-179062/16.

PT New artificial promoter libraries - containing consensus promoter

PT sequences and variable spacers, used to generate promoters for optimising

PS expression of genes.

PS Claim 28; Page 45; 89pp; English.

CC This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds
(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-10

Perfect score: 60
Sequence: 1 CATGACGAGTTATCTTCTG.....TATATATAAACACTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:
1: gb_ba:*
2: gb_beg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	60	100.0	60	BD006941	Artificia
2	21	35.0	2636	BC0081648	Danio rer
3	21	35.0	170386	BX248510	Zebrafish
4	20	33.3	60	BD006970	Artificia
5	20	31.7	653	BV051765	Artificia
6	19	31.7	4136	AF269996	Staphyloc
7	19	31.7	4136	AR485950	Sequence
8	19	31.7	4136	AX145314	Sequence
9	19	31.7	201628	AC101768	Mus muscu
10	19	31.7	223051	AC105554	Rattus no
11	19	31.7	288003	AC102971	Rattus no
12	19	31.7	303211	AB016746	Staphyloc
13	18	30.0	60	BD006936	Artificia
14	18	30.0	57200	AC151354	Xenopus t
15	18	30.0	59043	AC100881	Mus muscu
16	18	30.0	65708	AC151342	Xenopus t
17	18	30.0	86267	AC139048	Mus muscu
18	18	30.0	108277	AP006108	Locust cor
19	18	30.0	110000	BX908758_4	Continuation (5 of

20	18	30.0	118359	8	AP003415	Oriza sat
21	18	30.0	129336	9	AP005229	Homo sapi
22	18	30.0	143723	2	AC101763	Mus muscu
23	18	30.0	145726	2	AC021807	Homo sapi
24	18	30.0	147543	9	AC068189	Homo sapi
25	18	30.0	152103	2	AC025075	Homo sapi
26	18	30.0	153388	8	AP003289	Oriza sat
27	18	30.0	158146	9	AC005076	Homo sapi
28	18	30.0	161021	10	BX537253	Mouse DNA
29	18	30.0	163888	10	AC133185	Mus muscu
30	18	30.0	164928	9	AC087752	Homo sapi
31	18	30.0	164974	2	AC019082	Homo sapi
32	18	30.0	165129	9	AP003552	Homo sapi
33	18	30.0	166995	10	AC127229	Mus muscu
34	18	30.0	167762	5	BX511159	Zebrafish
35	18	30.0	168338	2	AC143421	Mus muscu
36	18	30.0	168617	10	AC125312	Mus muscu
37	18	30.0	175107	10	AC119911	Mus muscu
38	18	30.0	175403	9	AC093268	Homo sapi
39	18	30.0	176857	10	AC147184	Mus muscu
40	18	30.0	178208	2	AC136432	Homo sapi
41	18	30.0	179497	9	AL355580	Human DNA
42	18	30.0	179612	10	AC121970	Mus muscu
43	18	30.0	179711	2	AC019312	Homo sapi
44	18	30.0	181175	2	AC099623	Mus muscu
45	18	30.0	181927	9	AC114980	Homo sapi

ALIGNMENTS

RESULT 1
BD006941
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION
BD006941.1 GI:18635312
VERSION
JP 2001503249-A/10.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Lactococcus
REFERENCE
1 (bases 1 to 60)
AUTHORS
Hammer, K. and Janssen, P.R.
TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 10 13-MAR-2001;
PETER RUDAR JANSSEN
COMMENT.
OS Lactococcus lactis
PN JP 2001503249-A/10
PD 13-MAR-2001 JP 1998510287
PR 25-AUG-1997 JP 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
FH Key
FT promoter
Location/Qualifiers
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"
ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred No. 1.8e-22;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CATGACGAGTTATCTTCTGACACAGGTATGACTTATGATATATAAACACTACTGTT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-10

Perfect score: 60

Sequence: 1 CATGACGAGATTATTCTTG.....TATATATAAACACTACTGTT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_est7:
9: gb_est8:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	35.0	1006	9	CL117746 ISB1-69B1
2	19	31.7	216	4	B1678717 SW
3	19	31.7	477	5	BP702682 BP702682
4	19	31.7	546	7	CO251562 CO251562
5	19	31.7	584	7	CO251498 CO251498
6	19	31.7	718	8	BH294425 BH294425
7	19	31.7	793	7	CO488529 CO488529
8	19	31.7	808	7	CO226039 CO226039
9	19	31.7	911	9	CL110622 CL110622
10	19	31.7	921	9	CL135919 ISB1-107E
11	19	31.7	961	9	CL120712 ISB1-79F2
12	19	31.7	989	9	CL134757 ISB1-105H
13	19	31.7	991	9	CL134600 ISB1-105D
14	19	31.7	1024	9	CL021020 CH216-7M1
15	19	31.7	1078	9	CL082045 CH216-166
16	18	30.0	185	8	AZ815008 AZ815008
17	18	30.0	279	4	BM034608 BM034608
18	18	30.0	335	6	CD212844 CD212844
19	18	30.0	448	8	BF426602 BF426602
20	18	30.0	454	8	AZ123340 AZ123340
21	18	30.0	497	6	CD524954 CD524954
22	18	30.0	505	7	CF545817 CF545817
23	18	30.0	533	7	CF483680 POLI_24_A
24	18	30.0	535	9	CL326108 RBC14_24

25	18	30.0	564	8	BH341385 BH341385
26	18	30.0	580	8	AZ297021 RPCT-23-1
27	18	30.0	586	8	AZ025535 RPCT-23-3
28	18	30.0	589	7	CK383111 1a16e03.
29	18	30.0	618	7	CF430736 NIT1_2_A0
30	18	30.0	623	8	AZ824015 AZ824015
31	18	30.0	639	8	AZ387470 AZ387470
32	18	30.0	644	8	AQ977431 AQ977431
33	18	30.0	650	8	CR085296 CR085296
34	18	30.0	674	8	B2936668 B2936668
35	18	30.0	696	8	B79927 B79927
36	18	30.0	707	4	BJ643863 BJ643863
37	18	30.0	747	9	AG408050 AG408050
38	18	30.0	750	9	AG348322 AG348322
39	18	30.0	755	9	CL132595 CL132595
40	18	30.0	777	9	AG433214 AG433214
41	18	30.0	913	6	CA988543 CA988543
42	18	30.0	920	9	CNS02F49 CNS02F49
43	18	30.0	928	9	CL128193 CL128193
44	18	30.0	955	9	CL038779 CL038779
45	18	30.0	975	9	CL078264 CH216-148

ALIGNMENTS

RESULT 1
LOCUS CL117746 1006 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-69B10 Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-69B10,
genomic survey sequence.
ACCESSION CL117746
VERSION CL117746.1 GI:40611381
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 1006)
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E., and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 75000 Std Error: 0.00
Seg primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 4
High quality sequence stop: 617.
Location/Qualifiers
1..1006
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone_lib="ISB1-69B10"
/clone_1ib="ISB1"
/note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

FEATURES
source
1..1006
Location/Qualifiers

ORIGIN

Query Match 35.0%; Score 21; DB 9; Length 1006;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 TATGATATATATAAACACTGAC 56
DB 166 TATGATATATATAAACACTGAC 186

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-11

Perfect score: 60
Sequence: 1 CATGACNTAGTTCCTG.....TATATATATACGACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	60	100.0	60 2 AAV23183	AAV23183 Lactococcus
2	17	28.3	3123 6 ABA94996	Ab94996 Z. mays Z
3	16	26.7	256 5 ABV50605	Abv50605 Human pro
4	16	26.7	270 10 ADC74821	Adc74821 Human pro
5	16	26.7	580 8 AB254734	Ab254734 Aspergill
6	16	26.7	585 5 ABV47709	Abv47709 Human pro
7	16	26.7	3692 4 AAS01150	Aas01150 Interfero
8	16	26.7	3704 4 AAS01152	Aas01152 Interfero
9	16	26.7	5340 4 ABL06126	Ab106126 Drosophil
10	16	26.7	100445 13 ABD3179	Abd3179 Human can
11	16	26.7	186510 10 ADE24797	Ad24797 Human end
12	16	26.7	199878 10 ADL13719	Adl13719 Osteocarth
13	16	26.7	322885 13 ADS93537	Ads93537 Human MRC
14	15	25.0	245 4 AAC85419	Aac85419 Salmo sal
15	15	25.0	431 3 AAC28723	Aac28723 Human sec
16	15	25.0	555 13 ACN57955	Acn57955 Cotton gy
17	15	25.0	563 13 ACN49616	Acn49616 Cotton pr
18	15	25.0	728 3 AAA15002	Aaa15002 cDNA enco
19	15	25.0	876 9 ADA29158	Ada29158 DNA enco
20	15	25.0	1287 13 ADR85563	Adr85563 Aspergill

21	15	25.0	1352 13 ADR84976	Adr84976 Aspergill
22	15	25.0	2759 4 ABL25068	Ab125068 Drosophil
23	15	25.0	3140 4 ABL24165	Ab124165 Drosophil
24	15	25.0	3140 13 ADQ89679	Adq89679 Antagonis
25	15	25.0	7352 13 ADR84389	Adr84389 Aspergill
26	15	25.0	15644 5 AAS34611	Aas34611 Human DNA
27	15	25.0	17794 4 ABL02220	Ab102220 Drosophil
28	15	25.0	22693 11 ADF75185	Adf75185 Human NRG
29	15	25.0	27118 4 ABA08223	Ab08223 Human ova
30	15	25.0	48680 11 ACN45210	Acn45210 Human ova
31	15	25.0	52640 12 ADQ97220	Adq97220 Mouse can
32	15	25.0	96597 10 ADC85340	Adc85340 Human Lmo
33	15	25.0	96598 9 ADA02861	Ada02861 Mouse Itp
34	15	25.0	96598 10 ADM72599	Adm72599 Mouse Itp
35	15	25.0	96598 12 ADM74456	Adm74456 Murine ca
36	15	25.0	110000 3 AAP22305_08	Continuation (9 of
37	15	25.0	110000 6 ABA90521_09	Continuation (10 o
38	15	25.0	150130 11 ACN44394	Acn44394 Human gen
39	15	25.0	254366 8 AB223704	Ab223704 Human pho
40	15	25.0	326014 6 ABK89296	Abk89296 Human gen
41	15	25.0	326014 12 ADQ94981	Adq94981 Human kin
42	14	23.3	25 9 AC167074	Ac167074 Human mic
43	14	23.3	27 6 ABN88900	Abn88900 Human FGF
44	14	23.3	27 12 ADF86083	Adf86083 Mutagenic
45	14	23.3		

ALIGNMENTS

RESULT 1

ID AAV23183 standard; DNA; 60 BP.

XX AAV23183;

DT 28-JUL-1998 (first entry)

XX Lactococcus lactis constitutional promoter Cp15.

XX Lactococcus lactis; constitutional promoter; optimise; spacer;
artificial promoter library; gene expression; ds.

OS Synthetic.
OS Lactococcus lactis.

XX Key Location/Qualifiers
FT promoter 4..60
FT /*tag= a
FT /strand_name= "Constitutional promoter"

XX W09807846-A1.

XX 26-FEB-1998.

XX 25-AUG-1997; 97WO-DK000342.

XX 23-AUG-1996; 96DX-00000886.

XX (JENS/J) JENSEN P R.

XX Jensen PR; Hammer K;

XX WPI; 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter
sequences and variable spacers, used to generate promoters for optimising
expression of genes.

XX Claim 28; Page 46; 89pp; English.

XX This is a Lactococcus lactis constitutional promoter sequence used in the
construction of an artificial promoter library of the invention. The

Copyright 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-11

Perfect score: 60
Sequence: 1 CATTACNTAGTATCTCTG.....TATATATATACACTACTCTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Genemdb1:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006942	BD006942 Artificialia
2	18	30.0	100625	9 H51188X21	AL030997 Human DNA
3	18	30.0	165049	9 AL159995	AL159995 Human DNA
4	18	30.0	211511	10 AC125049	AC125049 Mus muscu
5	17	28.3	1167	1 D86380	D86380 Bacillus ce
6	17	28.3	3123	6 AX365269	AX365269 Sequence
7	17	28.3	4620	8 AF287472	AF287472 Arabidops
8	17	28.3	6507	1 AB041948	AB041948 Bacillus
9	17	28.3	8026	14 AY634314	AY634314 Solenopst
10	17	28.3	39838	2 AC149329	AC149329 Phakoposr
11	17	28.3	41700	9 AC108140	AC108140 Homo sapi
12	17	28.3	67077	2 AC100172	AC100172 Mus muscu
13	17	28.3	77720	8 ATT12C14	ATT12C14 Arabidops
14	17	28.3	97237	2 AP001949	AP001949 Homo sapi
15	17	28.3	105405	2 AC013325	AC013325 Homo sapi
16	17	28.3	120971	2 AC124966	AC124966 Medicago
17	17	28.3	143392	8 AC149299	AC149299 Populus b
18	17	28.3	153736	9 AC079301	AC079301 Homo sapi
19	17	28.3	154555	8 AC135958	AC135958 Oryza sat

C 20	17	28.3	156657	10 AL772253	AL772253 Mouse DNA
C 21	17	28.3	166377	2 AC102193	AC102193 Mus muscu
C 22	17	28.3	168595	9 AC013401	AC013401 Homo sapi
C 23	17	28.3	170533	2 AC105368	AC105368 Sus scrofa
C 24	17	28.3	173808	2 AC036206	AC036206 Homo sapi
C 25	17	28.3	177028	9 AC025043	AC025043 Homo sapi
C 26	17	28.3	179092	2 AC126949	AC126949 Homo sapi
C 27	17	28.3	190369	10 AC119908	AC119908 Rattus no
C 28	17	28.3	191236	9 AC090519	AC090519 Mus muscu
C 29	17	28.3	199257	10 AL928545	AL928545 Homo sapi
C 30	17	28.3	199385	9 AC092632	AC092632 Homo sapi
C 31	17	28.3	199774	2 AC095826	AC095826 Rattus no
C 32	17	28.3	212115	2 AC116701	AC116701 Mus muscu
C 33	17	28.3	228335	2 AC087875	AC087875 Mus muscu
C 34	17	28.3	240993	2 AC132771	AC132771 Rattus no
C 35	17	28.3	273994	2 CR792438	CR792438 Danio rer
C 36	17	28.3	317767	2 AC114654	AC114654 Mus muscu
C 37	16	26.7	256	6 C0518757	C0518757 Sequence
C 38	16	26.7	354	6 PB1251876	PB1251876 Phyllumed
C 39	16	26.7	585	6 C0515861	C0515861 Sequence
C 40	16	26.7	2238	10 BC083893	BC083893 Rattus no
C 41	16	26.7	2378	1 S5116945	S5116945 Staphylococ
C 42	16	26.7	3148	4 BTY08459	BTY08459 B. taurus mR
C 43	16	26.7	3480	1 SM040620	SM040620 Streptococ
C 44	16	26.7	3592	6 AX098234	AX098234 Sequence
C 45	16	26.7	3704	6 AX098238	AX098238 Sequence

ALIGNMENTS

RESULT 1
BD006942
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.

ACCESSION
BD006942.1 GI:18635313
VERSION
JP 2001503249-A/11.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Lactococcus lactis

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 60)
Hammer, K. and Janssen, P.R.
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
Patent: JP 2001503249-A 11 13-MAR-2001;
PETER RUDAR JANSSEN

OS Lactococcus lactis
PN JP 2001503249-A/11
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PI 23-AUG-1996 DK 0886/96
PR KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
Topology: Linear;
Key
FH promoter
FT Location/Qualifiers
1..60
Location/Qualifiers
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0% Score 60; DB 6; Length 60;
Best Local Similarity 100.0% Pred. No. 7.7e-25;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATTACNTAGTATCTCTGACGAGATTACGATTCGCTGATATATATACACTACTCTT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-11

Perfect score: 60
Sequence: 1 CATTACACTACTACTCTG.....TATATATATCACTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	19	31.7	157	5	BU086538
2	18	30.0	482	8	CR348683
3	18	30.0	503	8	AZ142620
4	18	30.0	982	9	CNSO4D75
5	17	28.3	288	5	BM895253
6	17	28.3	359	8	AZ140839
7	17	28.3	467	8	AQ177575
8	17	28.3	609	8	BZ915314
9	17	28.3	669	2	BB629352
10	17	28.3	718	9	CE360435
11	17	28.3	772	8	CC308769
12	17	28.3	775	9	CR301188
13	17	28.3	803	8	BH589927
14	17	28.3	845	9	CG899772
15	17	28.3	854	9	CG365987
16	17	28.3	896	9	CG365976
17	17	28.3	1090	3	AK036450
18	17	28.3	1091	9	CL081071
19	17	28.3	1260	3	AY108249
20	16	26.7	209	5	BM300214
21	16	26.7	256	8	BZ292667
22	16	26.7	271	1	AA447698
23	16	26.7	300	7	CF143628
24	16	26.7	304	8	CC424690

c 25	16	26.7	308	8	AA019587
c 26	16	26.7	364	5	BY021270
c 27	16	26.7	365	8	CC393858
c 28	16	26.7	388	8	AQ252517
c 29	16	26.7	423	9	CG480349
c 30	16	26.7	424	8	AG205257
c 31	16	26.7	430	2	AM630985
c 32	16	26.7	440	4	BG511574
c 33	16	26.7	440	8	AZ751199
c 34	16	26.7	474	6	CD186225
c 35	16	26.7	477	6	CD186246
c 36	16	26.7	494	8	CC164315
c 37	16	26.7	508	9	CE805695
c 38	16	26.7	514	8	AZ563958
c 39	16	26.7	518	7	CF080570
c 40	16	26.7	553	2	AM654247
c 41	16	26.7	556	2	AM654254
c 42	16	26.7	557	9	CE270237
c 43	16	26.7	567	8	A0699007
c 44	16	26.7	588	9	CE363558
c 45	16	26.7	601	5	BX085873

ALIGNMENTS

RESULT 1
BU086538
LOCUS
DEFINITION Na L3 09E06 SAC Necator americanus (parasitic nematode) L3 Necator americanus cDNA clone Na_L3_09E06 5', mRNA sequence.
ACCESSION BU086538.1 GI:22527727
VERSION
KEYWORDS
SOURCE
ORGANISM
Necator americanus
Necator americanus
Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Strongylida; Ancylostomatoidae; Ancylostomatidae; Bunostominae; Necator.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared for Dr. David Pritchard University of Nottingham. Sequencing was performed by the Pathogen Sequencing Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail and Bart Barrell). The sequence contained a Polya tail (trimmed)
PCR Primers
FORWARD: SAC
BACKWARD: T7PL
Plate: 09 row: E column: 06
Seg primer: SAC
High quality sequence stop: 157.

FEATURES

source
1..157
/organism="Necator americanus"
/mol_type="mRNA"
/db_xref="taxon:51031"
/clone="Na_L3_09E06"
/sex="mixed"
/dev_stage="L3"
/clone_lib="Necator americanus (parasitic nematode) L3"
/note="Vector: PCMV-PCR vector; Site 1: EcoRI (5' end); Site 2: XhoI (3' end); Necator americanus is a human hookworm, responsible for debilitating anaemia. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 129.623 Seconds

(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-12

Perfect score: 58
Sequence: 1 CATGTGTAGTTATCTCTG.....GGTATATATACAGTACTGAG 58

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003ds:*\n11: geneseqn2004as:*\n12: geneseqn2004bs:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	58	100.0	AAV23184	AAV23184 Lactococc
2	19	32.8	AAV23203	AAV23203 Lactococc
3	18	31.0	AAV23207	AAV23207 Lactococc
4	18	31.0	AAV23197	AAV23197 Lactococc
5	17	29.3	AAV23209	AAV23209 Lactococc
6	17	29.3	AAV23212	AAV23212 Lactococc
7	17	29.3	AAV23178	AAV23178 Lactococc
8	17	29.3	AAV23178	AAV23178 Lactococc
9	17	29.3	AAV23178	AAV23178 Lactococc
10	17	29.3	AAV23178	AAV23178 Lactococc
11	17	29.3	AAV23178	AAV23178 Lactococc
12	17	29.3	AAV23178	AAV23178 Lactococc
13	17	29.3	AAV23178	AAV23178 Lactococc
14	17	29.3	AAV23178	AAV23178 Lactococc
15	17	29.3	AAV23178	AAV23178 Lactococc
16	17	29.3	AAV23178	AAV23178 Lactococc
17	17	29.3	AAV23178	AAV23178 Lactococc
18	17	29.3	AAV23178	AAV23178 Lactococc
19	17	29.3	AAV23178	AAV23178 Lactococc
20	17	29.3	AAV23178	AAV23178 Lactococc

21	16	27.6	60	2	AAV23188	AAV23188 Lactococc
22	16	27.6	60	2	AAV23191	AAV23191 Lactococc
23	16	27.6	60	2	AAV23189	AAV23189 Lactococc
24	16	27.6	60	2	AAV23213	AAV23213 Lactococc
25	16	27.6	60	2	AAV23214	AAV23214 Lactococc
26	16	27.6	60	2	AAV23204	AAV23204 Lactococc
27	16	27.6	60	2	AAV23198	AAV23198 Lactococc
28	16	27.6	60	2	AAV23205	AAV23205 Lactococc
29	16	27.6	64	6	ABK98604	ABK98604 L. lactis
30	16	27.6	64	9	ACD13855	ACD13855 L. lactis
31	16	27.6	65	6	ABK98605	ABK98605 L. lactis
32	16	27.6	65	9	ACD13856	ACD13856 L. lactis
33	16	27.6	93	6	ABK98615	ABK98615 L. lactis
34	16	27.6	93	9	ACD13866	ACD13866 L. lactis
35	16	27.6	207	6	ABK98627	ABK98627 L. lactis
36	16	27.6	207	9	ACD13878	ACD13878 L. lactis
37	16	27.6	242	6	ABK98617	ABK98617 L. lactis
38	16	27.6	303	6	ABL77210	ABL77210 Human ova
39	16	27.6	303	6	ABL77210	ABL77210 Human ova
40	16	27.6	374	4	AA184069	AA184069 Human pol
41	16	27.6	399	8	ABX54585	ABX54585 Bovine ES
42	16	27.6	414	3	AAH30481	AAH30481 Human col
43	16	27.6	422	4	AAK57283	AAK57283 Human lmm
44	16	27.6	422	8	ABX47777	ABX47777 Bovine ES
45	16	27.6	437	6	ABQ73760	ABQ73760 Human col

ALIGNMENTS

RESULT 1
ID AAV23184 standard; DNA; 58 BP.

AAV23184;

28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cp16.

Lactococcus lactis; constitutional promoter; optimise; spacer;
artificial promoter library; gene expression; ds.

Synthetic.
Lactococcus lactis.

Key Location/Qualifiers

Promoter

4.58
/*tag= a
/standard_name= "Constitutional promoter"

W09807846-A1.

26-FEB-1998.

25-AUG-1997; 97MO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENSEN) JENSEN P R.

Jensen PR, Hammer K;

WPL; 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 46; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

1 CATGTGTAGTTATCTTGACAGCATGAGTCAATTTGGTATATAACAGTACTCAG 58

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1185.61 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-12

Perfect score: 58

Sequence: 1 CATTGCTACTTATCTTG.....GGTATATACAGTCTCAG 58

Scoring table: OLIGO_NUC

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	19	32.8	667 5	B0634768 NXR073.A
2	19	32.8	1125 8	CC223944 CH261.250
3	18	31.0	286 1	AV062189 AV062189
4	18	31.0	293 8	AO675730 HS_5489.B
5	18	31.0	296 1	AV099749 AV099749
6	18	31.0	484 6	CD083525 MA3-9999U
7	18	31.0	509 9	CNS00UYX
8	18	31.0	526 9	CNS00WOD
9	18	31.0	549 8	AO962707 Arabidops
10	18	31.0	558 8	BH745962 LERGH20TF
11	18	31.0	570 8	AO969890 LERUR11TR
12	18	31.0	571 8	AO962833 LERGH77TF
13	18	31.0	583 8	AO962832 LERGH77TF
14	18	31.0	584 8	AO961635 LERGA44TR
15	18	31.0	587 1	AV517995 AV517995
16	18	31.0	589 8	AO961634 LERGA44TR
17	18	31.0	590 8	BH008110 eel7e03.X
18	18	31.0	595 8	AO969889 LERUR11TR
19	18	31.0	604 9	CE639628 tigr-988-
20	18	31.0	612 6	CA387595 663608.NC
21	18	31.0	622 8	BH010041 egl1e09.X
22	18	31.0	626 9	CC959251 BOID178TF
23	18	31.0	640 8	B2033232 oel188a11.
24	18	31.0	652 6	CB259306 63-B9622-

25	18	31.0	657 8	BH743231	BH743231 g27a10.9
26	18	31.0	657 8	AG081932	AG081932 Pan trogl
27	18	31.0	665 8	BH947149	BH947149 cbv1h12.
28	18	31.0	666 8	BH743464	BH743464 g27a09.9
29	18	31.0	668 8	BH996241	BH996241 oel89a02.
30	18	31.0	673 8	AO956467	AO956467 LERAX17TF
31	18	31.0	673 8	BZ055350	BZ055350 jnr29g02.
32	18	31.0	674 9	CC957239	CC957239 BOIBR33TR
33	18	31.0	676 8	BH960317	BH960317 oed98c05.
34	18	31.0	676 8	BH996463	BH996463 oej27h05.
35	18	31.0	677 8	BH922912	BH922912 odg85d11.
36	18	31.0	684 8	BZ014043	BZ014043 oel17g03.
37	18	31.0	684 8	BZ075005	BZ075005 lke35a06.
38	18	31.0	686 8	BH921167	BH921167 odj26d12.
39	18	31.0	686 8	BH967602	BH967602 odg74d04.
40	18	31.0	690 8	BH993879	BH993879 oel28h10.
41	18	31.0	692 8	BZ020051	BZ020051 oel13d03.
42	18	31.0	693 8	BH958553	BH958553 odj30a12.
43	18	31.0	696 8	BZ048418	BZ048418 jnr55b11.
44	18	31.0	696 8	BZ048471	BZ048471 jnr55b11.
45	18	31.0	696 8	BZ063734	BZ063734 lke72b05.

ALIGNMENTS

RESULT 1
LOCUS B0634768 667 bp mRNA linear EST 07-MAY-2003
DEFINITION NXR073 A06 F NXRV (Nsf Xylem Root wood Vertical) Pinus taeda cDNA
clone NXR073_A06 5', mRNA sequence.
B0634768
B0634768.1 GI:21688921

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pinus taeda (loblolly pine)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 667)
Sederoff, R.
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Sederoff, Ron
Forest Biotechnology
North Carolina State University
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
NC 27695, USA
Tel: 919 515 7800
Fax: 919 515 7801
Email: ron_sederoff@ncsu.edu, jerry_johnson@ncsu.edu
Please see http://web.dnc.unm.edu/biodata/nsfpine/ for further
information.
Seq primer: T3.

FEATURES

source

Location/Qualifiers
1..667
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXRV073 A06"
/cdate="19970731"
/cdate="19970731"
/cell_type="Root (primary)"
/dev_stage="Transitional"
/lab_host="XLA-Blue"
/clone_lib="NXRV (Nsf Xylem Root wood Vertical)"
/note="Vector: pBluescript SK-; Site 1: Eco RI; Site 2:
XhoI. The library is from primary xylem scraped from the
roots of a twelve year old tree in the transitional phase
from juvenile wood to mature wood production. NOTE: The
sequences contain a 'cDNA adapter' between the Score1 site
and the start of the EST. The adapter sequence is
'AATCGGACGACG'."

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 131.858 Seconds
(Without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657b-13

Perfect score: 59
Sequence: 1 CATCTCGAGTATTATCTTG.....TATTAATAGTACGACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	59	2	AAV23185
2	17	28.8	60	2	AAV23200
3	17	28.8	1887	10	ADBS6066
4	17	28.8	23075	5	ABAI9112
5	17	28.8	38239	12	ADG97626
6	16	27.1	59	2	AAV23179
7	16	27.1	59	2	AAV23210
8	16	27.1	60	2	AAV23211
9	16	27.1	60	2	AAV23177
10	16	27.1	60	2	AAV23191
11	16	27.1	60	2	AAV23182
12	16	27.1	60	2	AAV23199
13	16	27.1	236	3	AAV23182
14	16	27.1	415	13	ADG55705
15	16	27.1	450	10	ADG55705
16	16	27.1	450	10	ADG55705
17	16	27.1	450	10	ADG55705
18	16	27.1	450	10	ADG55705
19	16	27.1	450	10	ADG55705
20	16	27.1	450	10	ADG55705

21	16	27.1	504	2	AAV23185
22	16	27.1	539	13	ADG58054
23	16	27.1	579	9	ACH26916
24	16	27.1	585	4	AAH12182
25	16	27.1	627	3	AAH1787
26	16	27.1	775	2	AAH18595
27	16	27.1	782	2	AAH18595
28	16	27.1	909	2	AAH18595
29	16	27.1	1198	2	AAH18595
30	16	27.1	1378	4	AAH18595
31	16	27.1	2304	2	AAH18595
32	16	27.1	2307	2	AAH18595
33	16	27.1	2368	8	ADG55705
34	16	27.1	2368	10	ADG55705
35	16	27.1	2368	13	ADG55705
36	16	27.1	2567	12	ADG55705
37	16	27.1	2567	12	ADG55705
38	16	27.1	2791	5	ABAI9112
39	16	27.1	3611	3	AAH18595
40	16	27.1	3780	8	ADG55705
41	16	27.1	3780	10	ADG55705
42	16	27.1	3780	13	ADG55705
43	16	27.1	6690	5	ADG55705
44	16	27.1	8927	5	ADG55705
45	16	27.1	11740	13	ADG55705

ALIGNMENTS

RESULT 1	AAV23185	standard; DNA; 59 BP.
XX	AAV23185	
XX	AAV23185	
AC	AAV23185	
XX	AAV23185	
DT	28-JUL-1998	(first entry)
XX	AAV23185	
DE	Lactococcus lactis constitutional promoter Cp17.	
XX	Lactococcus lactis; constitutional promoter; optimise; spacer;	
KM	artificial promoter library; gene expression; de.	
XX	Synthetic.	
OS	Lactococcus lactis.	
XX	Lactococcus lactis.	
FT	Key	Location/Qualifiers
FT	promoter	4..59
FT	promoter	/*tag= a
FT	promoter	/standard_name= "Constitutional promoter"
XX	WO9807846-A1.	
XX	26-FEB-1998.	
PD	25-AUG-1997;	97WO-DK000342.
PF	25-AUG-1997;	97WO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
PR	23-AUG-1996;	96DK-00000886.
XX	(JENS/) JENSEN P R.	
PA	(JENS/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
XX	Jensen PR, Hammer K;	
DR	WPI, 1998-179062/16.	
XX	WPI, 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX	Claim 28; Page 47; 89pp; English.	
PS	This is a Lactococcus lactis constitutional promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 458.474 Seconds
(without alignments)

6235.594 Million cell updates/sec

Title: US-09-242-657B-13

Perfect score: 1 CATCTCGAGTTATTCTTG.....TATTAATGATACAGTCTGT 59

Sequence:

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank1:
1: gb Da:
2: gb Htg:
3: gb In:
4: gb Om:
5: gb Ov:
6: gb Pat:
7: gb Ph:
8: gb Pl:
9: gb Pr:
10: gb Ro:
11: gb Sts:
12: gb Sy:
13: gb Un:
14: gb Vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	BD006944	BD006944 Artificialia
2	21	35.6	247958	AC108340	AC108340 Rattus no
3	20	33.9	122122	AL139420	AL139420 Human DNA
4	18	30.5	177708	AC093804	AC093804 Homo sapi
5	18	30.5	178075	AC021615	AC021615 Homo sapi
6	7	30.5	182970	AL935328	AL935328 Mouse DNA
7	18	30.5	185371	AC083795	AC083795 Homo sapi
8	9	30.5	191070	AC134254	AC134254 Mus muscu
9	18	30.5	201389	AC093358	AC093358 Mus muscu
10	18	30.5	203794	AC140930	AC140930 Mus muscu
11	18	30.5	206200	AC093318	AC093318 Mus muscu
12	18	30.5	223779	AC120443	AC120443 Rattus no
13	18	30.5	347286	AC102740	AC102740 Mus muscu
14	17	28.8	60	BD006959	BD006959 Artificialia
15	17	28.8	191	BV079516	BV079516 D17C882
16	17	28.8	12412	AC006669	AC006669 Caenorhab
17	17	28.8	14951	AB000756	AB000756 Aquifex a
18	17	28.8	31228	U42841	U42841 Caenorhabdi
19	17	28.8	42596	U41263	U41263 Caenorhabdi

20	17	28.8	55237	8	CNS08CR6	AL928743 Oryza sat
C 21	17	28.8	60101	2	AC101124	AC101124 Mus muscu
C 22	17	28.8	84076	2	AC149797	AC149797 Aedes aeg
C 23	17	28.8	100000	9	AP000081	AP000081 Homo sapi
C 24	17	28.8	109222	8	CNS08C91	AL733378 Oryza sat
C 25	17	28.8	110000	8	AB016819	Continuation (15 o
C 26	17	28.8	115006	2	OSJN00110	AL606694 Oryza sat
C 27	17	28.8	123066	2	AC093105	AC093105 Schistosm
C 28	17	28.8	127683	9	AC003683	AC003683 Homo sapi
C 29	17	28.8	145366	2	AC024945	AC024945 Homo sapi
C 30	17	28.8	146074	2	AC141806	AC141806 Adis mell
C 31	17	28.8	146167	9	AC087588	AC087588 Homo sapi
C 32	17	28.8	146640	5	AC149787	AC149787 Gallus ga
C 33	17	28.8	147042	5	AL935296	AL935296 Xenopus t
C 34	17	28.8	148504	2	AC147906	AC147906 Human chr
C 35	17	28.8	149326	9	CNS01RHP	AL162471 Human chr
C 36	17	28.8	151104	2	AC115001	AC115001 Mus muscu
C 37	17	28.8	151479	8	AC093713	AC093713 Oryza sat
C 38	17	28.8	152511	2	AC019269	AC019269 Homo sapi
C 39	17	28.8	159473	10	BX247953	BX247953 Mouse DNA
C 40	17	28.8	162626	2	AC138551	AC138551 Dario rer
C 41	17	28.8	164452	9	CNS01RGP	AL159140 Human chr
C 42	17	28.8	166001	2	HS4462D18	AL132765 Human DNA
C 43	17	28.8	170497	2	AC117037	AC117037 Rattus no
C 44	17	28.8	172070	2	AC134951	AC134951 Dario rer
C 45	17	28.8	182314	10	AL935301	AL935301 Mouse DNA

ALIGNMENTS

RESULT 1
BD006944 59 bp DNA linear PAT 31-JAN-2002
LOCUS Artificial promoter libraries for selected organisms and promoters
DEFINITION derived from such libraries.

ACCESSION BD006944.1 GI:18635315
VERSION JP 2001503249-A/13.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE 1 (bases 1 to 59)
AUTHORS Hammer,K. and Janssen,P.R.
TITLE Artificial promoter libraries for selected organisms and promoters
JOURNAL derived from such libraries
Patent: JP 2001503249-A 13 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT OS Lactococcus lactis
PN JP 2001503249-A/13
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96

PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
FH Topology: Linear;
FT Key Location/Qualifiers
promoter (4)..(59).
Location/Qualifiers

FEATURES
source 1..59
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN
Query Match 100.0%; Score 59; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 8.7e-25;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATCTCGAGTTATTCTTGACCGCTGATGACAGTGGTATTAATGATGACTGTT 59
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1206.06 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-13

Perfect score: 59
Sequence: 1 CATTCTGAGTATTACTTG.....TATATAGTACACTGCTT 59

Scoring table: OLIGO_NTC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	33.9	912	8	BZ401648
C 2	20	33.9	1145	4	EM013229
C 3	18	30.5	330	7	CN951547
C 4	18	30.5	362	2	AM945946
C 5	18	30.5	468	2	AM845451
C 6	18	30.5	817	9	CG768533
C 7	17	28.8	206	5	BU431614
C 8	17	28.8	359	2	AM127434
C 9	17	28.8	360	1	AV201509
C 10	17	28.8	360	7	D34913
C 11	17	28.8	394	5	BY371615
C 12	17	28.8	403	8	AO738202
C 13	17	28.8	405	5	BY371805
C 14	17	28.8	443	9	CR108885
C 15	17	28.8	449	8	AO726829
C 16	17	28.8	517	1	AA304949
C 17	17	28.8	586	1	AO799774
C 18	17	28.8	602	5	BM056650
C 19	17	28.8	603	5	BM056650
C 20	17	28.8	612	5	BM322630
C 21	17	28.8	628	5	BM370478
C 22	17	28.8	637	5	BM363995
C 23	17	28.8	653	5	BM359662
C 24	17	28.8	671	4	BI270176

25	17	28.8	710	8	BH943899	BH943899	odf17h08.
C 26	17	28.8	712	9	CE557431	CE557431	cigr-gs-
27	17	28.8	760	9	CC904398	CC904398	t04f18ba
28	17	28.8	767	9	AG575700	AG575700	mus muscu
29	17	28.8	778	9	CC484887	CC484887	CH240_314
C 30	17	28.8	790	9	CC863504	CC863504	NDL_41P16
C 31	17	28.8	800	5	BU281873	BU281873	603601774
C 32	17	28.8	862	8	BH210249	BH210249	PUGCH72TB
C 33	17	28.8	892	8	AZ542327	AZ542327	ENTR82TR
C 34	17	28.8	913	8	AZ690972	AZ690972	ENTR02TR
C 35	17	28.8	913	8	BH165847	BH165847	ENTRM35TF
C 36	17	28.8	915	8	AZ546153	AZ546153	ENTRM87TR
C 37	17	28.8	931	8	BZ810251	BZ810251	PUGCH72TD
C 38	17	28.8	954	8	BH130939	BH130939	ENTRM19TF
C 39	17	28.8	1210	8	CC281159	CC281159	CH261-76K
C 40	16	27.1	101	7	W49714	W49714	zc43e12.r1
C 41	16	27.1	159	7	N50575	N50575	yy89b04.r1
C 42	16	27.1	188	8	AQ103607	AQ103607	HS_3069_B
43	16	27.1	196	1	AA808838	AA808838	nm47h10.r
44	16	27.1	199	1	AI810355	AI810355	wb87f03.x
45	16	27.1	199	1	AI972587	AI972587	wf40e02.x

ALIGNMENTS

RESULT 1
BZ401648/c 912 bp DNA linear GSS 04-DEC-2002
LOCUS
DEFINITION
OGART95TC ZM 0.7 1.5 KB Zea mays genomic clone ZMMEMa0025B03,
genomic survey sequence.

ACCESSION
BZ401648
VERSION
BZ401648.1 GI:26026718

KEYWORDS
GSS.

SOURCE
ORGANISM
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta, Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 912)

White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Cleck, R.W., Numborg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics
Unpublished (2002)

Other GSSs: OGART95TM
Contact: Cathy White,law

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white@tigr.org
Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..912
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMEMa0025B03"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source

ORIGIN

Query Match 33.9%; Score 20; DB 8; Length 912;
Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 GATATATAGTACACTGCT 58
DB 840 GATATATAGTACACTGCT 821

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 129.623 Seconds

(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-14

Perfect score: 58
Sequence: 1 CATTTGCACTTATCTTG.....GTATTAATCACTACTGTT 58

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	100.0	58	2	AAV23186
2	19	32.8	59	2	AAV23206
3	19	32.8	60	2	AAV23194
4	18	31.0	59	2	AAV23193
5	18	31.0	60	6	ABK98622
6	18	31.0	60	6	ABK98622
7	18	31.0	95	6	ABK98612
8	18	31.0	95	6	ABK98612
9	18	31.0	9437	6	ABK98601
10	18	31.0	9437	9	ACD13852
11	18	31.0	10929	6	ABK98591
12	18	31.0	10929	6	ACD13842
13	18	31.0	12739	6	ABK98592
14	18	31.0	12739	6	ACD13843
15	17	29.3	3524	4	AB126502
16	17	29.3	160771	6	ABO88179
17	16	27.6	58	6	ABK98625
18	16	27.6	58	9	ACD13876
19	16	27.6	59	2	AAV23210
20	16	27.6	60	2	AAV23188

21	16	27.6	60	2	AAV23200	AAV23200 Lactococc
22	16	27.6	60	2	AAV23212	AAV23212 Lactococc
23	16	27.6	60	2	AAV23199	AAV23199 Lactococc
24	16	27.6	64	6	ABK98604	ABK98604 L. lactis
25	16	27.6	64	9	ACD13855	ACD13855 L. lactis
26	16	27.6	65	6	ABK98605	ABK98605 L. lactis
27	16	27.6	65	6	ABK98605	ABK98605 L. lactis
28	16	27.6	93	6	ABK98615	ABK98615 L. lactis
29	16	27.6	93	6	ABK98615	ABK98615 L. lactis
30	16	27.6	115	6	ABK98626	ABK98626 L. lactis
31	16	27.6	115	9	ACD13877	ACD13877 L. lactis
32	16	27.6	120	12	ADL97740	ADL97740 Human tra
33	16	27.6	150	6	ABK98616	ABK98616 L. lactis
34	16	27.6	150	9	ACD13867	ACD13867 L. lactis
35	16	27.6	431	5	ABK15226	ABK15226 Human ner
36	16	27.6	431	5	ABK15226	ABK15226 Human ner
37	16	27.6	477	3	ACG37168	ACG37168 Arabidops
38	16	27.6	755	6	ABN99221	ABN99221 Arabidops
39	16	27.6	1062	6	ABZ13652	ABZ13652 Arabidops
40	16	27.6	1062	6	ADG87654	ADG87654 A. thalia
41	16	27.6	1062	6	ADG87654	ADG87654 A. thalia
42	16	27.6	1062	8	ADA68053	ADA68053 Arabidops
43	16	27.6	1305	3	ACG47953	ACG47953 Arabidops
44	16	27.6	1305	3	ACG51562	ACG51562 Arabidops
45	16	27.6	2118	3	AAK51464	AAK51464 Arabidops

ALIGNMENTS

RESULT 1	AAV23186	standard; DNA; 58 BP.
ID	AAV23186	
AC	AAV23186	
XX	28-JUL-1998	(first entry)
DE	Lactococcus lactis	constitutional promoter Cp18.
KW	Lactococcus lactis	constitutional promoter; optimise; spacer;
XX	artificial promoter library; gene expression; ds.	
OS	Synthetic.	
XX	Lactococcus lactis.	
FT	Key	Location/Qualifiers
FT	promoter	4..58
XX	/*tag= a	
XX	/standard_name= "Constitutional promoter"	
XX	W09807846-A1.	
XX	26-FEB-1998.	
XX	25-AUG-1997;	97MO-DR000342.
XX	23-AUG-1996;	96DK-00000886.
XX	(JENS/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
XX	WPI; 1998-119062/16.	
DR	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX	Claim 28; Page 47; 89pp; English.	
CC	This is a Lactococcus lactis constitutional promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 450.703 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657b-14

Perfect score: 58
Sequence: 1 CATTTCGAGTTTATCTCTG.....GTATAACTACTACTGTT 58Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0
Total number of hits satisfying chosen parameters: 9416466Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenBml:*
1: gb ba:*
2: gb htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	58	100.0	58	6 BD006945	BD006945 Artificia
2	58	34.5	186406	10 AL713394	AL713394 Mouse DNA
3	19	32.8	59	6 BD006965	BD006965 Artificia
4	19	32.8	60	6 BD006963	BD006963 Artificia
5	19	32.8	59387	2 AC100266	AC100266 Mus muscu
6	19	32.8	128294	8 AC135225	AC135225 Oryza sat
7	19	32.8	137580	8 AF377947	AF377947 Oryza sat
8	19	32.8	157970	8 AC092556	AC092556 Oryza sat
9	19	32.8	204367	10 AC115055	AC115055 Mus muscu
10	19	32.8	210334	2 AC128609	AC128609 Rattus no
11	19	32.8	228178	2 AC095775	AC095775 Rattus no
12	19	32.8	251079	2 AC134743	AC134743 Rattus no
13	19	31.0	59	6 BD006952	BD006952 Artificia
14	19	31.0	331	11 G35155	G35155 crase Crypt
15	19	31.0	1087	8 AK071640	AK071640 Oryza sat
16	19	31.0	3003	8 AK066362	AK066362 Oryza sat
17	19	31.0	4643	8 AK068491	AK068491 Oryza sat
18	19	31.0	32412	3 CEF53B6	CE53B6 Caenorhadi
19	19	31.0	93029	9 AL354941	AL354941 Human DNA

20	18	31.0	100986	8 AC105363	AC105363 Oryza sat
21	18	31.0	106806	5 AL935319	AL935319 Zebrafish
22	18	31.0	126258	8 AP005758	AP005758 Oryza sat
23	18	31.0	126827	8 AC118980	AC118980 Oryza sat
24	18	31.0	149527	2 CR391909	CR391909 Danio rer
25	18	31.0	154019	10 AL671885	AL671885 Mouse DNA
26	18	31.0	159792	8 AP005894	AP005894 Oryza sat
27	18	31.0	163150	2 AL158161	AL158161 Homo sapi
28	18	31.0	174447	2 AC142552	AC142552 Danio rer
29	18	31.0	177104	10 AC091237	AC091237 Mus muscu
30	18	31.0	182160	2 AC142550	AC142550 Danio rer
31	18	31.0	184810	10 AC090869	AC090869 Mus muscu
32	18	31.0	190790	10 AC102582	AC102582 Mus muscu
33	18	31.0	190850	2 AC015833	AC015833 Homo sapi
34	18	31.0	191431	2 AC148740	AC148740 Oryctolag
35	18	31.0	195778	5 BX088566	BX088566 Zebrafish
36	18	31.0	200704	9 AC099781	AC099781 Homo sapi
37	18	31.0	219216	2 AC115506	AC115506 Rattus no
38	18	31.0	226494	2 AC106498	AC106498 Rattus no
39	18	31.0	227039	2 AC123395	AC123395 Rattus no
40	18	31.0	227565	2 AC112761	AC112761 Rattus no
41	18	31.0	231453	2 CR847527	CR847527 Danio rer
42	18	31.0	234584	2 AC094320	AC094320 Rattus no
43	18	31.0	234801	2 AC118309	AC118309 Rattus no
44	18	31.0	250099	2 AC093949	AC093949 Rattus no
45	18	31.0	253236	2 BX950188	BX950188 Danio rer

ALIGNMENTS

RESULT 1	BD006945	58 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006945				
DEFINITION	Artificial promoter libraries for selected organisms and promoters derived from such libraries.				
ACCESSION	BD006945	1	GI:18635316		
VERSION	BD006945.1				
KEYWORDS	UP 2001503249-A/14.				
SOURCE	Lactococcus lactis				
ORGANISM	Lactococcus lactis				
REFERENCE	1 (bases 1 to 58) Hammer,K. and Janssen,P.R. Artificial promoter libraries for selected organisms and promoters derived from such libraries Patent: JP 2001503249-A 14 13-MAR-2001;				
JOURNAL	PETER RUDAR JANSSEN				
COMMENT	OS Lactococcus lactis				
	PN UP 2001503249-A/14				
	PD 13-MAR-2001				
	PF 25-AUG-1997 JP 1998510287				
	PR 23-AUG-1996 DK 0886/96				
	PI KALIN HAMMER, PETER RUDAR JANSSEN				
	PC C12N15/09, C12N15/00				
	CC Strandedness: Double;				
	CC Topology: Linear;				
	FH Key				
	FT promoter				
FEATURES	Location/Qualifiers (4)..(58).				
source	1..58				
	/organism="Lactococcus lactis"				
	/mol_type="Genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN	100.0%; Score 58; DB 6; Length 58;				
Query Match	Best Local Similarity 100.0%; Pred. No. 5.4e-22;				
Matches	58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
1	CATTTCGAGTTTATCTCTGACATGTGTGCGGTGATTAATACTACTACTGTT 58				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1185.61 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-14

Perfect score: 58
Sequence: 1 CATTTCGACGTTTCTCTG.....GTATTACTACTACTGTT 58

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	34.5	815	8	CC097349 CSU-K34.1
2	19	32.8	236	8	BH739830 GF20412.B
3	19	32.8	647	8	AQ666076 HS 5358.B
4	19	32.8	671	9	CE597453 tigr-gss-
5	19	32.8	721	9	AG372426 Mus muscu
6	19	32.8	757	9	AG372386 Mus muscu
7	19	32.8	772	9	AG477703 Mus muscu
8	19	32.8	825	9	CL940941 OA ABA005
9	18	31.0	437	8	B43798 HS-1058-A2-
10	18	31.0	437	8	NS8914 YF612.2.s1
11	18	31.0	438	8	AZ199120 SP 1039.A
12	18	31.0	479	8	AJ685409 AJ685409
13	18	31.0	494	6	CD054360 H001K23r
14	18	31.0	590	6	CA263833 SCRLB203
15	18	31.0	685	6	CE345490 tigr-gss-
16	18	31.0	709	7	CF303069 ABF1-01-
17	18	31.0	844	6	CB676481 OSJNB12M
18	18	31.0	919	6	CB663351 OSJNB12M
19	18	31.0	923	6	CB635589 OSJNB12M
20	18	31.0	971	9	AG378261 Mus muscu
21	18	31.0	1111	8	CC277811 CH261-30K
22	18	31.0	1662	6	CG755518 P051-1-B0
23	17	29.3	155	6	CD988241 OAP106.Y
24	17	29.3	261	2	BF402972 UR-R-CA0-

25	17	29.3	278	9	CC776736	CC776736 ZMBRC042
26	17	29.3	289	7	CR471509	CR471509 CR471509
27	17	29.3	321	7	CR627270	CR627270 000303AAL
28	17	29.3	349	9	CG681691	CG681691 ZMBRC016
29	17	29.3	352	5	BU549582	BU549582 GM880024B
30	17	29.3	394	8	AO181303	AO181303 HS 3223.A
31	17	29.3	396	2	BR870199	BR870199 IL3-BT011
32	17	29.3	407	5	BU545490	BU545490 GM880005B
33	17	29.3	411	4	BG726429	BG726429 sad44101.
34	17	29.3	414	9	CC527321	CC527321 CH240.402
35	17	29.3	431	2	AM202386	AM202386 sf14d07.Y
36	17	29.3	449	2	AW119960	AW119960 LjNEST12E
37	17	29.3	453	2	CG282314	CG282314 OG4AH12TC
38	17	29.3	459	2	BF401576	BF401576 UR-R-CA0-
39	17	29.3	460	7	CN862162	CN862162 000817AAL
40	17	29.3	470	6	CD995688	CD995688 QB3a04.Y
41	17	29.3	526	4	BI290656	BI290656 UR-R-DKO-
42	17	29.3	527	8	BZ970934	BZ970934 PUF1A65TD
43	17	29.3	548	7	CF922457	CF922457 gmrHRw24
44	17	29.3	548	8	BZ848371	BZ848371 CH240_252
45	17	29.3	557	5	BU551327	BU551327 GM880023B

ALIGNMENTS

RESULT 1
LOCUS CC097349 815 bp DNA linear GSS 16-APR-2003
DEFINITION CSU-K34.107K3.SP6 CSU-K34 Aedes aegypti genomic clone
ACCESSION CC097349
VERSION CC097349.1 GI:29961795
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes; Stegomyia.
REFERENCE 1 (bases 1 to 815)
Loftus, B., Shetty, J., Severson, D., Brown, S. and Knudson, D.
End sequencing of Aedes aegypti BACs
Unpublished (2003)
Other GSSs: CSU-K34.107K3.T7
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by Susan Brown and Dennis Knudson at Colorado State University.
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..815
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/db_xref="taxon:7159"
/clone="CSU-K34.107K3"
/clone_id="CSU-K34"
/note="Vector: pBAC3.6; Site 1: EcoRI; Source DNA: Aedes aegypti, strain unknown [derived from freshly hatched larvae at the Virus Research Centre, Poona, India.
Reference: SINGH, K. R. P., 1967 Cell cultures derived from larvae of Aedes albopictus (Skuse) and Aedes aegypti (L.). Current Science 36: 506-508; ATC-10 cell line ATCC CCL-125"

ORIGIN
Query Match 34.5%; Score 20; DB 8; Length 815;
Best Local Similarity 100.0%; Pred. No. 1.7;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 129.623 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-15

Perfect score: 58

Sequence: 1 CATGCTTAGTTCCTTCTGA.....ATTATATAGTACTACTGTT 58

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 16Dec04:*

1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	58	2	AAV23187 Lactococc
2	17	29.3	19480	4	AAK80384 Human imm
3	17	29.3	19481	4	AAK80383 Human imm
4	17	29.3	42123	11	ACN44906 Human gen
5	17	29.3	107036	13	ABD33557 Human ost
6	17	29.3	147708	6	ABQ88154 Human can
7	17	29.3	158001	12	ADL17884 Human pho
8	16	27.6	530	13	ACN48042 Cotton pr
9	16	27.6	564	6	ABN62784 Human can
10	16	27.6	1003	2	AAV52511 Streptoco
11	16	27.6	1317	4	AAV52511 Streptoco
12	16	27.6	3020	13	ADSB0593 Bacterial
13	16	27.6	3447	10	ADSB0593 Bacterial
14	16	27.6	3504	4	AAQ02483 C. neofor
15	16	27.6	33454	11	ACN44588 Mouse gen
16	16	27.6	34688	6	ABO67059 Human ang
17	16	27.6	45315	13	ABD33576 Murine ca
18	16	27.6	110000	10	ABSS56454_11 Continuation (12 o
19	16	27.6	218336	8	ABQ76678 Androgen
20	16	27.6	266145	10	ADB87477 Fowlpox v

ALIGNMENTS

21	16	27.6	34980	5	AAH41225	Aah41225 Pyrococc
22	15	25.9	60	2	AAV23177	Aav23177 Lactococc
23	15	25.9	190	6	ABL78998	Abk178998 Human ova
24	15	25.9	339	6	ABK81005	Abk81005 Bacillus
25	15	25.9	339	6	ABK80981	Abk80981 Bacillus
26	15	25.9	339	6	ABV18899	Abv18899 Human pro
27	15	25.9	395	10	ABR41741	Abt41741 Toxicity
28	15	25.9	397	10	ACH29819	Ach29819 Human tes
29	15	25.9	400	4	AAH72448	Aah72448 Human cer
30	15	25.9	423	3	ACH45941	Aach45941 Human inf
31	15	25.9	434	3	AAAC98810	Aac98810 Human pan
32	15	25.9	467	5	ABV47920	Abv47920 Human pro
33	15	25.9	478	5	ABV48679	Abv48679 Human pro
34	15	25.9	499	4	AAH69555	Aah69555 Human cer
35	15	25.9	500	12	ACH74641	Ach74641 Human gen
36	15	25.9	500	12	ACH74116	Ach74116 Human gen
37	15	25.9	510	12	ACH75849	Ach75849 Human gen
38	15	25.9	519	6	ABO56279	Abg56279 Human ova
39	15	25.9	524	5	AD169170	Adi69170 Human ova
40	15	25.9	524	5	AD175517	Adi75517 Human ova
41	15	25.9	530	5	AAO08710	Aad08710 Lycoperetl
42	15	25.9	556	6	ABV97061	Abv97061 Human pan
43	15	25.9	567	6	AAV70125	AAV70125 DNA encod
44	15	25.9	567	6	ABQ23172	Abq23172 Oligonuct
45	15	25.9	567	6	ABQ23173	Abq23173 Oligonuct

RESULT 1

ID AAV23187 standard; DNA; 58 BP.

AC AAV23187;

DT 28-JUL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cp19.

KM Lactococcus lactis; constitutional promoter; optimise; spacer;

XX artificial promoter library; gene expression; de.

OS Synthetic.

XX Lactococcus lactis.

FT Key Location/Qualifiers

FT promoter 4..58 /tag= a

XX /standard_name= "Constitutional promoter"

XX WO9807846-A1.

XX 26-FEB-1998.

XX 25-AUG-1997; 97WO-DK000342.

XX 23-AUG-1996; 96DK-00000886.

XX (JENS/) JENSEN P R.

XX Jensen PR, Hammer K;

XX WPI, 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter

XX sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.

XX Claim 28; Page 48; 89pp; English.

XX This is a Lactococcus lactis constitutional promoter sequence used in the

XX construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 450.703 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657b-15

Perfect score: 58
Sequence: 1 CATCGCTTACTTTTCTTGA.....ATTATATAGTTAGTACTGTT 58

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	58	100.0	58	BD006946	BD006946 Artificial
2	58	32.8	199386	AC131738	AC131738 Mus muscu
3	18	31.0	695990	AC087771	AC087771 Genomic S
4	18	31.0	143048	CR751542	CR751542 Danio rer
5	18	31.0	195552	CR405690	CR405690 Zebrafish
6	18	31.0	224957	AC114173	AC114173 Rattus no
7	17	29.3	358	AF217361	AF217361 Helicater
8	17	29.3	992	BC081402	BC081402 Danio rer
9	17	29.3	2667	AF127135	AF127135 Homo sapi
10	17	29.3	7030	AB048278	AB048278 Mus muscu
11	17	29.3	20062	AC017789	AC017789 Drosophila
12	17	29.3	37297	AF047657	AF047657 Genomhab
13	17	29.3	65718	H0311G09	H0311G09 Oryza sat
14	17	29.3	84833	AC149532	AC149532 Xenopus t
15	17	29.3	99271	HS808P6	HS808P6 Human DNA
16	17	29.3	110000	AB017282	AB017282_24
17	17	29.3	110000	AP006840	AP006840_30
18	17	29.3	113462	AC150205	Continuation (31 o
19	17	29.3	119304	AP004255	AP004255 Oryza sat

c	20	17	29.3	135727	9	AC005999	AC005999 Homo sapi
c	21	17	29.3	138487	8	AP005875	AP005875 Oryza sat
c	22	17	29.3	139376	9	AC020659	AC020659 Homo sapi
c	23	17	29.3	141894	5	AL627129	AL627129 Zebrafish
c	24	17	29.3	142023	8	AP004705	AP004705 Oryza sat
c	25	17	29.3	143263	8	AC130724	AC130724 Oryza sat
c	26	17	29.3	144177	2	AL513487	AL513487 Human DNA
c	27	17	29.3	146080	2	AC136858	AC136858 Rattus no
c	28	17	29.3	147708	9	HS179M20	HS179M20 Human DNA
c	29	17	29.3	148937	8	AP004815	AP004815 Oryza sat
c	30	17	29.3	150116	2	AC034132	AC034132 Homo sapi
c	31	17	29.3	153502	2	AC063925	AC063925 Homo sapi
c	32	17	29.3	155895	5	BX470221	BX470221 Zebrafish
c	33	17	29.3	157654	2	AC091791	AC091791 Sus scrofa
c	34	17	29.3	158498	10	AL807801	AL807801 Mouse DNA
c	35	17	29.3	160867	9	AC025254	AC025254 Homo sapi
c	36	17	29.3	160931	3	AC093498	AC093498 Drosophila
c	37	17	29.3	162445	9	AL158151	AL158151 Human DNA
c	38	17	29.3	163932	2	AC146625	AC146625 Papio anu
c	39	17	29.3	164823	8	AP005007	AP005007 Oryza sat
c	40	17	29.3	165797	2	BX649400	BX649400 Danio rer
c	41	17	29.3	168103	2	AC149435	AC149435 Pan trogl
c	42	17	29.3	168413	2	AC009877	AC009877 Homo sapi
c	43	17	29.3	168458	9	AC019278	AC019278 Homo sapi
c	44	17	29.3	169305	9	AC021036	AC021036 Homo sapi
c	45	17	29.3	172482	9	AC022309	AC022309 Homo sapi

ALIGNMENTS

RESULT 1	BD006946	58 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006946				
DEFINITION	Artificial promoter libraries for selected organisms and promoters derived from such libraries.				
ACCESSION	BD006946				
VERSION	BD006946.1				GI:18635317
KEYWORDS	JP 2001503249-A/15.				
SOURCE	Lactococcus lactis				
ORGANISM	Lactococcus lactis				
REFERENCE	Lactococcus. 1 (bases 1 to 58)				
AUTHORS	Hammer, K. and Janssen, P. R.				
TITLE	Artificial promoter libraries for selected organisms and promoters derived from such libraries				
JOURNAL	Patent: JP 2001503249-A 15 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
OS	Lactococcus				
PN	JP 2001503249-A/15				
PD	13-MAR-2001				
PF	25-AUG-1997				JP 1998510287
PR	23-AUG-1996				DK 0886/96
PI	KALIN HAMMER, PETER RUDAR JANSSEN				
PC	CI2N15/09, CI2N15/00				
CC	Strandedness: Double;				
CC	Topology: Linear;				
FT	Key				Location/Qualifiers
FT	promoter				(4)..(58).
FEATURES	Location/Qualifiers				
source	1..58				
	/organism="Lactococcus lactis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN	Query Match 100.0%; Score 58; DB 6; Length 58;				
	Best Local Similarity 100.0%; Pred. No. 3e-23;				
	Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 CATCGCTTACTTTTCTTGAAGAGGATCGGGTGTATATATAGTACTGTT 58				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1185.61 Seconds
(Without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-15

Perfect score: 58
Sequence: 1 CATCGCTACTTTTCTTGA.....ATATAATAGTACTACTGTT 58

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_ests1:
2: gb_ests2:
3: gb_ests3:
4: gb_ests4:
5: gb_ests5:
6: gb_ests6:
7: gb_ests7:
8: gb_ests8:
9: gb_ests9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	DB ID	Description
1	21	36.2	1216	4	BT259996 602972095
2	19	32.8	271	4	BM300609
3	19	32.8	678	9	CC631567 OGUBR76TH
4	19	32.8	698	8	CC170867 1168905.9
5	19	32.8	754	9	CG033969 PUKA042TD
6	19	32.8	817	8	CC335010 OGUNAT2TV
7	19	32.8	885	9	CG033966 PUKA042TB
8	19	32.8	929	9	CG048627 PUFUJ37TD
9	19	32.8	1118	2	BE037088 MP14F05 M
10	19	31.0	487	7	CK104680 1050P05.5
11	18	31.0	570	7	CK501421 ramga0_00
12	18	31.0	577	7	CK500601 rambo_01
13	18	31.0	609	8	BH076366 RPT-24-2
14	18	31.0	625	7	CK552605 rsw1a0_00
15	18	31.0	813	1	AV055885 AV055885
16	18	31.0	896	7	CK021997 ACENCCURT
17	18	31.0	957	5	BUI144379 603228131
18	18	31.0	1114	6	BZ692100 SP_Ba001
19	17	29.3	271	6	CD450161 USDA-FP-1
20	17	29.3	305	4	BT553555
21	17	29.3	321	7	CK687782 ZF101-P00
22	17	29.3	457	7	CV149846 Mact6007D
23	17	29.3	482	7	CF305953 HPA1--02-
24	17	29.3	579	8	BZ337450 1867n05.9

c	25	17	29.3	601	8	BZ302264	BZ302264	KD1740.P1
c	26	17	29.3	604	5	B0778843	946114F09	
c	27	17	29.3	612	8	AZ366675	1M0116F06	
c	28	17	29.3	625	8	AQ576990	nbxb0090K	
c	29	17	29.3	638	9	CL834290	OR_CBA005	
c	30	17	29.3	640	7	CO353327	DR_ATE_FL	
c	31	17	29.3	689	9	BX126108	Danio Ter	
c	32	17	29.3	695	7	CK6727143	CRK727143	ZF101-P00
c	33	17	29.3	704	9	CR226795	Reverse s	
c	34	17	29.3	723	7	CO960149	AGENCOURT	
c	35	17	29.3	726	9	CC699626	OGVFP66TV	
c	36	17	29.3	728	8	BZ67251	PUDG14TD	
c	37	17	29.3	739	7	CO795256	AGENCOURT	
c	38	17	29.3	741	9	CG351216	CG28822TH	
c	39	17	29.3	742	9	CG351222	OG28822TV	
c	40	17	29.3	756	8	BZ041217	11r20b02.	
c	41	17	29.3	757	7	CN015220	AGENCOURT	
c	42	17	29.3	759	9	CC669007	OGMB223TH	
c	43	17	29.3	772	8	A0869968	nbeb0036G	
c	44	17	29.3	776	9	CR070604	Reverse s	
c	45	17	29.3	800	9	CG285008	OG0E157TV	

ALIGNMENTS

RESULT 1
BT259996 1216 bp mRNA linear EST 17-JUL-2001
LOCUS 602972095F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:511615 5',
DEFINITION RNA sequence.
ACCESSION BT259996
VERSION BT259996.1 GI:14817875
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1216)
NIH-MGC http://mgs.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E.B. Consortium/ILNL at:
http://image.llnl.gov
Plate: ILNL1270 row: m column: 24
High quality sequence stop: 162.
Location/Qualifiers
1..1216
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:511615"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 36.2%; Score 21; DB 4; Length 1216;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 TTTCTTACAGAGAGATCC 32

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-16

Perfect score: 60
Sequence: 1 CATTTGCTAGTTATCTTG.....TATATCTAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2 AAV23188	AAV23188 Lactococc
2	19	31.7	60	2 AAV23212	AAV23212 Lactococc
3	17	28.3	59	2 AAV23193	AAV23193 Lactococc
4	17	28.3	60	2 AAV23214	AAV23214 Lactococc
5	17	28.3	60	6 ABK98622	ABK98622 Lambda CP
6	17	28.3	60	6 ABK98612	ABK98612 Lambda CP
7	17	28.3	95	6 ABK98612	ABK98612 Lambda CP
8	17	28.3	95	6 ABK98612	ABK98612 Lambda CP
9	17	28.3	207	6 ABK98627	ABK98627 L. lactis
10	17	28.3	207	6 ABK98627	ABK98627 L. lactis
11	17	28.3	242	6 ABK98617	ABK98617 L. lactis
12	17	28.3	242	6 ABK98617	ABK98617 L. lactis
13	17	28.3	9437	6 ABK98601	ABK98601 Vector pE
14	17	28.3	9437	6 ABK98601	ABK98601 Vector pE
15	17	28.3	9566	6 ABK98634	ABK98634 Vector pE
16	17	28.3	9566	6 ABK98634	ABK98634 Vector pE
17	17	28.3	10929	6 ABK98591	ABK98591 pEP23 Vec
18	17	28.3	10929	6 ABK98591	ABK98591 pEP23 Vec
19	17	28.3	12739	6 ABK98592	ABK98592 Vector pE
20	17	28.3	12739	9 ACD13843	ACD13843 Plasmid p

21	17	28.3	110000	6 ABA90521_00	ABA90521 Genomic s
22	17	28.3	185371	6 ABT10718	ABT10718 Human bre
23	17	28.3	193672	10 ADI13570	ADI13570 Osteoartr
24	16	26.7	58	2 AAV23186	AAV23186 Lactococc
25	16	26.7	58	2 AAV23184	AAV23184 Lactococc
26	16	26.7	59	2 AAV23209	AAV23209 Lactococc
27	16	26.7	59	2 AAV23210	AAV23210 Lactococc
28	16	26.7	59	2 AAV23206	AAV23206 Lactococc
29	16	26.7	60	2 AAV23194	AAV23194 Lactococc
30	16	26.7	60	2 AAV23197	AAV23197 Lactococc
31	16	26.7	60	2 AAV23213	AAV23213 Lactococc
32	16	26.7	60	2 AAV23200	AAV23200 Lactococc
33	16	26.7	60	2 AAV23178	AAV23178 Lactococc
34	16	26.7	60	2 AAV23199	AAV23199 Lactococc
35	16	26.7	60	2 AAV23205	AAV23205 Lactococc
36	16	26.7	282	10 ABK84336	ABK84336 Corn ear-
37	16	26.7	505	6 ABS61417	ABS61417 Prostate
38	16	26.7	516	13 ADQ55158	ADQ55158 Novel can
39	16	26.7	797	3 AAC49974	AAC49974 Arabidops
40	16	26.7	799	3 AAC40360	AAC40360 Arabidops
41	16	26.7	974	4 AAK84726	AAK84726 Human Imm
42	16	26.7	2306	4 ABL25306	ABL25306 Drosophi
43	16	26.7	2369	4 ABL25308	ABL25308 Drosophi
44	16	26.7	45839	6 AAD42191	AA42191 Human ngy
45	16	26.7	107330	12 ADQ97316	ADQ97316 Mouse can

ALIGNMENTS

RESULT 1

ID AAV23188 standard; DNA; 60 BP.

AC AAV23188;

DT 28-JUL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cp2.

KW Lactococcus lactis; constitutional promoter; optimise; spacer;

XX Lactococcus lactis.

OS Synthetic.

OS Lactococcus lactis.

FT Key

FT promoter

PN WO9807846-A1.

PD 26-FEB-1998.

XX 25-AUG-1997; 97WO-DK000342.

PR 23-AUG-1996; 96DK-00000886.

XX (JENS/) JENSEN P R.

PI Jensen PR, Hammer K;

XX WPI; 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter

PT sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.

PS Claim 28; Page 48; 89pp; English.

XX This is a Lactococcus lactis constitutional promoter sequence used in the

CC construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-16

Sequence: 1 CATTGCTAGTATTACTCTTG.....TATATTACTGAGTACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Genembl:*

1: gb_ba:*\n2: gb_hrg:*\n3: gb_in:*\n4: gb_om:*\n5: gb_ov:*\n6: gb_pat:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_ro:*\n11: gb_scs:*\n12: gb_sy:*\n13: gb_un:*\n14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006947	BD006947 Artificia
2	19	31.7	60	BD006971	BD006971 Artificia
3	19	31.7	206016	AC133808	AC133808 Rattus no
4	19	31.7	224789	AC111611	AC111611 Rattus no
5	19	31.7	262424	AC103010	AC103010 Rattus no
6	18	30.0	72909	AL833795	AL833795 Mouse DNA
7	18	30.0	90157	AL355552	AL355552 Human DNA
8	18	30.0	101753	AC116125	AC116125 Mus muscu
9	18	30.0	106953	AP004188	AP004188 Oryza sat
10	18	30.0	107830	HS146B4	HS146B4 Homo sapi
11	18	30.0	126263	AP004650	AP004650 Oryza sat
12	18	30.0	135330	AL772165	AL772165 Mouse DNA
13	18	30.0	145503	5 BX470238	5 BX470238 Zebrafish
14	18	30.0	146073	9 HS147D16	9 HS147D16 Homo sapi
15	18	30.0	153070	2 AL355361	2 AL355361 Human DNA
16	18	30.0	154677	9 AC016196	9 AC016196 Homo sapi
17	18	30.0	185801	10 AC113127	10 AC113127 Mus muscu
18	18	30.0	188687	2 BX640405	2 BX640405 Danio rer
19	18	30.0	191672	2 AC020794	2 AC020794 Mus muscu

20	18	30.0	219218	2 AC020817	AC020817 Mus muscu
21	18	30.0	250029	9 AE014310	AE014310 Homo sapi
22	18	30.0	260573	2 AC125775	AC125775 Rattus no
23	18	30.0	340000	2 HS21C084	HS21C084 Homo sapi
24	17	28.3	59	6 BD006952	6 BD006952 Artificia
25	17	28.3	60	6 BD006973	6 BD006973 Artificia
26	17	28.3	67	6 AR055663	6 AR055663 Sequence
27	17	28.3	815	5 CR353147	5 CR353147 Gallus ga
28	17	28.3	1100	1 LITRRRM	1 LITRRRM Lactis ge
29	17	28.3	1100	6 AR055662	6 AR055662 Sequence
30	17	28.3	6850	8 AY193555	8 AY193555 Lactuca s
31	17	28.3	12737	1 AE006241	1 AE006241 Lactococc
32	17	28.3	36019	9 AC108726	9 AC108726 Homo sapi
33	17	28.3	62513	2 AC139686	2 AC139686 Homo sapi
34	17	28.3	65942	2 AC036124	2 AC036124 Homo sapi
35	17	28.3	68301	2 AL158065	2 AL158065 Human DNA
36	17	28.3	73021	2 AC017368	2 AC017368 Drosophila
37	17	28.3	76587	9 AL390954	9 AL390954 Human DNA
38	17	28.3	106214	8 AP003818	8 AP003818 Oryza sat
39	17	28.3	109381	9 AP002350	9 AP002350 Homo sapi
40	17	28.3	130973	8 AP003813	8 AP003813 Oryza sat
41	17	28.3	133659	10 AC092094	10 AC092094 Mus muscu
42	17	28.3	136814	9 AC064826	9 AC064826 Homo sapi
43	17	28.3	139123	9 AL513527	9 AL513527 Human DNA
44	17	28.3	139695	5 BX324128	5 BX324128 Zebrafish
45	17	28.3	140730	9 AC010942	9 AC010942 Homo sapi

ALIGNMENTS

RESULT 1
BD006947 60 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.

ACCESSION
BD006947
VERSION
BD006947.1 GI:18635318
KEYWORDS
JP 2001503249-A/16.
SOURCE
Lactococcus lactis
ORGANISM
Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
1 (bases 1 to 60)
AUTHORS
Hammer K. and Janssen P.R.
TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 16 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT
OS
Lactococcus lactis
PN
JP 2001503249-A/16
PD
13-MAR-2001
PF
25-AUG-1997 JP 1998510287
PR
23-AUG-1996 DK 0886/96
PI
KALIN HAMMER, PETER RUDAR JANSSEN
PC
C12N15/09, C12N15/00
CC
Strandedness: Double;
Topology: Linear;
FH
Key
promoter
Location/Qualifiers
(4). (60).
Location/Qualifiers
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1558"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1,1e-21;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATTGCTAGTATTACTCTTGACATGACGCGCTTATGATGATTTACTGACTGTT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-16

Perfect score: 60
Sequence: 1 CATTCGCTAGTTTACTTCTTG.....TATATCTAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST:

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est1.*
9: gb_est2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	268	6	CD166614 MM1-0005P
2	19	31.7	383	7	CO980069 GMB9007A1
3	19	31.7	1289	9	CG752793 P047-3-FO
4	18	30.0	1206	9	CL504131 SATL 733
5	17	28.3	208	7	CF473502 RTW2 3 D
6	17	28.3	224	1	BB074525 BB074525
7	17	28.3	324	2	AU091720 AU091720
8	17	28.3	329	5	BT305856 BT305856
9	17	28.3	351	5	BY009356 BY009356
10	17	28.3	353	9	CR199258 CR199258
11	17	28.3	379	8	AZ304492 AZ304492
12	17	28.3	419	8	AQ431624 AQ431624
13	17	28.3	445	9	CL361142 CL361142
14	17	28.3	489	9	TA86004P TA86004P
15	17	28.3	510	8	AQ972629 RPCT-23-3
16	17	28.3	537	6	CB280227 CB280227
17	17	28.3	570	9	CL385164 CL385164
18	17	28.3	572	1	AJ724350 AJ724350
19	17	28.3	601	4	BJ313250 BJ313250
20	17	28.3	623	7	CN048077 CN048077
21	17	28.3	631	9	AG245251 AG245251
22	17	28.3	655	2	AM695272 NF093D11S
23	17	28.3	661	8	BZ334983 BZ334983
24	17	28.3	678	8	AZ274153 RPCT-23-1

C 25	17	28.3	683	7	CN625364	CN625364 tae01a11.
C 26	17	28.3	703	7	CF764101	CF764101 C8E004548
C 27	17	28.3	707	8	AZ977561	AZ977561 2M0253F05
C 28	17	28.3	727	9	CL454251	CL454251 ZMMBBD048
C 29	17	28.3	746	8	CC110884	CC110884 NDL 70012
C 30	17	28.3	763	5	BU435712	BU435712 603221095
C 31	17	28.3	812	9	CR803938	CR803938 GR00A24D
C 32	17	28.3	830	7	CV499470	CV499470 63829.1 M
C 33	17	28.3	847	8	AZ669235	AZ669235 EMTMS77P
C 34	17	28.3	849	8	AZ137692	AZ137692 SP 0174 B
C 35	17	28.3	856	9	CL057249	CL057249 CH216-85A
C 36	17	28.3	869	7	CO214789	CO214789 WS00930. B
C 37	17	28.3	1603	9	CG754186	CG754186 P049-3-H0
C 38	16	26.7	182	7	U94691	U94691 PVU94691 Pe
C 39	16	26.7	254	6	CA829448	CA829448 3529.1.5
C 40	16	26.7	261	9	CL891308	CL891308 abg07C02.
C 41	16	26.7	264	8	BZ724375	BZ724375 PUDNA36TB
C 42	16	26.7	267	6	CB466477	CB466477 731900 MA
C 43	16	26.7	286	8	AZ303845	AZ303845 IM0003D10
C 44	16	26.7	286	9	CE159286	CE159286 t1gr-988-
C 45	16	26.7	293	6	CD330702	CD330702 STRP537.

ALIGNMENTS

RESULT 1
CD166614 268 bp mRNA linear EST 14-SEP-2003
LOCUS
DEFINITION MM1-0005P-V006-B05-U-B MM1-0005 Schistosoma mansoni cDNA clone
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

CD166614.1 GI:34703258
EST.
Schistosoma mansoni
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 268)
Verjovski-Almeida, S., Demarco, R., Martins, E.A.L., Guimaraes, P.E.M.,
Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y., Jr.,
Kittajima, J.P., Adamson, R.B., Ashton, P.D., Bonaldo, M.F.,
Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,
Lente, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,
Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,
Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,
Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,
Setubal, J.C., Lente, L.C.C. and Dias-Neto, E.
Transcriptome analysis of the acclimated human parasite Schistosoma
mansoni
Nat. Genet. 35 (2), 148-157 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquímica
Instituto de Química - Universidade de São Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,
Brasil
Tel: +55-11-3091-2172
Fax: +55-11-3091-2186
Email: verjovski@iq.usp.br
This sequence was derived from the PAPERP Schistosoma mansoni EST
Genome Project. All sequences in the project were assembled and
annotated. This entry and all the assembled sequences can be seen
in the following URL: <http://biointo.iq.usp.br/schisto/>
Plate: MM1-0005P-V006 Row: 5 Column: B.

FEATURES
source
1..268
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MM1-0005P-V006-B05.B"
/sex="male"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-17

Perfect score: 60
Sequence: 1 CATGGGTGAGTTATCTTGT.....TATCATAGCAGACTACTAT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	AAV23189	AAV23189 Lactococc
2	19	31.7	444	ADs48396	Ades48396 Bacterial
3	18	30.0	60	AAV23177	AAV23177 Lactococc
4	18	30.0	60	AAV23198	AAV23198 Lactococc
5	17	28.3	59	AAV23179	AAV23179 Lactococc
6	17	28.3	60	AAV23191	AAV23191 Lactococc
7	17	28.3	60	AAV23204	AAV23204 Lactococc
8	17	28.3	265	AAQ23880	AAQ23880 pKTH1.820
9	17	28.3	265	AAQ46205	AAQ46205 Promoter
10	17	28.3	265	AAV23187	AAV23187 Promoter
11	17	28.3	2864	ADP50636	ADP50636 Murine DN
12	17	28.3	110000	AAV23184	AAV23184 Lactococc
13	16	26.7	58	AAV23184	AAV23184 Lactococc
14	16	26.7	58	ABK98625	ABK98625 L. lactis
15	16	26.7	58	ACD13876	ACD13876 L. lactis
16	16	26.7	59	AAV23190	AAV23190 Lactococc
17	16	26.7	59	AAV23210	AAV23210 Lactococc
18	16	26.7	60	AAV23211	AAV23211 Lactococc
19	16	26.7	60	AAV23203	AAV23203 Lactococc
20	16	26.7	60	AAV23207	AAV23207 Lactococc

ALIGNMENTS

21	16	26.7	60	2	AAV23182	AAV23182 Lactococc
22	16	26.7	60	2	AAV23200	AAV23200 Lactococc
23	16	26.7	60	2	AAV23178	AAV23178 Lactococc
24	16	26.7	60	2	AAV23199	AAV23199 Lactococc
25	16	26.7	64	6	ABK98604	ABK98604 L. lactis
26	16	26.7	64	6	ACD13855	ACD13855 L. lactis
27	16	26.7	65	6	ABK98605	ABK98605 L. lactis
28	16	26.7	65	6	ACD13856	ACD13856 L. lactis
29	16	26.7	93	6	ABK98615	ABK98615 L. lactis
30	16	26.7	93	6	ACD13866	ACD13866 L. lactis
31	16	26.7	456	12	ADP93380	ADP93380 Cotton ex
32	16	26.7	755	6	ABN99221	ABN99221 Arabidops
33	16	26.7	905	4	ABL11525	ABL11525 Drosophil
34	16	26.7	1062	6	ABZ13652	ABZ13652 Arabidops
35	16	26.7	1062	6	ADG87654	ADG87654 A. thalia
36	16	26.7	1062	6	ADG87655	ADG87655 A. thalia
37	16	26.7	1062	8	ADA68053	ADA68053 Arabidops
38	16	26.7	1305	3	ACA47953	ACA47953 Arabidops
39	16	26.7	1305	3	ACA51562	ACA51562 Arabidops
40	16	26.7	1473	4	ABL11531	ABL11531 Drosophil
41	16	26.7	2000	8	ADA72010	ADA72010 Rice gene
42	16	26.7	2118	3	ACA51464	ACA51464 Arabidops
43	16	26.7	2714	4	ABL21716	ABL21716 Drosophil
44	16	26.7	2845	4	ABL11728	ABL11728 Drosophil
45	16	26.7	3019	4	ABP09002	ABP09002 Drosophil

RESULT 1

ID	AAV23189	standard, DNA, 60 BP.
AC	AAV23189;	
DT	28-JUL-1998	(first entry)
DE	Lactococcus lactis constitutional promoter Cp20.	
KW	Lactococcus lactis; constitutional promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; ds.	
OS	Synthetic.	
OS	Lactococcus lactis.	
PH	Key	Location/Qualifiers
FT	Promoter	4..60
FT		/*tag= a
XX		/standard_name= "Constitutional promoter"
PN	W09807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97MO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENS/) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K,	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
PS	Claim 28, Page 49; 89pp; English.	
XX		
CC	This is a Lactococcus lactis constitutional promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005. CompuGen Ltd.

OM nucleic - nucleic search, using BW model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds
(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657b-17

Perfect score: 60
Sequence: 1 CATGGGTGAGTTTATCTCTG.....TATCATAGCAGAGTACTATT 60

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBml:*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	BD006948	Artificia
2	19	31.7	1962	AR420479	Trichobil
3	19	31.7	57455	ACT100895	Mus muscu
4	19	31.7	71024	NCB21D9	Neurospor
5	19	31.7	158867	AC067851	Homo sapi
6	19	31.7	159237	AC138328	Mus muscu
7	19	31.7	180525	AC107834	Mus muscu
8	19	31.7	187708	ACT108850	Mus muscu
9	19	31.7	188098	ACT108396	Mus muscu
10	19	31.7	206074	AC009484	Homo sapi
11	19	31.7	214848	AC100561	Mus muscu
12	19	31.7	224164	AC125310	Mus muscu
13	18	30.0	60	BD006936	Artificia
14	18	30.0	73170	BD006957	Artificia
15	18	30.0	73170	AL445068	Human DNA
16	18	30.0	110000	AC121143	Continuatio
17	18	30.0	175703	AC121143	Mus muscu
18	18	30.0	184892	AC016017	Mus muscu
19	18	30.0	194362	AC124716	Mus muscu

c	20	18	30.0	209643	9438 Mus muscu
c	21	18	30.0	239732	490 Rattus no
c	22	18	30.0	313163	243 Rattus no
c	23	17	28.3	59	938 Artificia
c	24	17	28.3	60	950 Artificia
c	25	17	28.3	60	963 Artificia
c	26	17	28.3	68	668 Sequence
c	27	17	28.3	258	6 L. lactis pl
c	28	17	28.3	265	AR363960
c	29	17	28.3	1470	AK119468
c	30	17	28.3	3378	AK099595
c	31	17	28.3	10029	AK099595
c	32	17	28.3	13040	AK099595
c	33	17	28.3	31570	AK099595
c	34	17	28.3	65392	AK099595
c	35	17	28.3	84821	AK099595
c	36	17	28.3	85624	AK099595
c	37	17	28.3	89473	AK099595
c	38	17	28.3	110000	AK099595
c	39	17	28.3	110000	AK099595
c	40	17	28.3	110000	AK099595
c	41	17	28.3	127259	AK099595
c	42	17	28.3	136263	AK099595
c	43	17	28.3	140980	AK099595
c	44	17	28.3	142649	AK099595
c	45	17	28.3	144404	AK099595

ALIGNMENTS

RESULT 1	BD006948	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006948	Artificial promoter libraries for selected organisms and promoters			
DEFINITION	BD006948	Artificial promoter libraries for selected organisms and promoters			
ACCESSION	BD006948.1	GI:18635319			
VERSION	JP 2001503249-A/17				
KEYWORDS	Lactococcus lactis				
SOURCE	Lactococcus lactis				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
REFERENCE	1 (bases 1 to 60)				
AUTHORS	Hammer, K. and Janssen, P.R.				
TITLE	Artificial promoter libraries for selected organisms and promoters				
JOURNAL	derived from such libraries				
COMMENT	Patent: JP 2001503249-A 17 13-MAR-2001; PETER RUDAR JANSSEN				
OS	Lactococcus lactis				
PN	JP 2001503249-A/17				
PD	13-MAR-2001				
PP	25-AUG-1997 JP 1998510287				
PR	23-AUG-1996 DK 0866/96				
PC	KALIN HAMMER, PETER RUDAR JANSSEN				
PI	C12N15/09, C12N15/00				
CC	Strandedness: Double;				
CC	Topology: Linear;				
CC	Key				
FT	promoter				
FEATURES	Location/Qualifiers				
source	1..60				
	/organism="Lactococcus lactis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN					
Query Match	100.0%; Score 60; DB 6; Length 60;				
Best Local Similarity	100.0%; Pred. No. 3.4e-25;				
Matches	60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 CATGGGTGAGTTTATCTCTGACGTCGCGCCGCGGCGCTATATCATAGCAGTACTATT 60				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-17

Sequence: 1 CATGGGTGAGTTATCTTCTG.....TATCATAGCAGTACTATT 60

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	30.0	310	1	AV299623 AV299623
2	18	30.0	459	1	BU787017 BU787017
3	18	30.0	1281	9	AG430622 AG430622
4	17	28.3	352	9	CG629520 CG629520
5	17	28.3	361	1	AU186307 AU186307
6	17	28.3	459	9	CE541862 CE541862
7	17	28.3	461	7	CO306209 CO306209
8	17	28.3	466	8	AQ188220 AQ188220
9	17	28.3	515	8	AQ962389 AQ962389
10	17	28.3	576	9	CT749099 CT749099
11	17	28.3	581	7	CF426773 CF426773
12	17	28.3	606	8	AQ962388 AQ962388
13	17	28.3	628	7	CF416219 CF416219
14	17	28.3	645	4	BM337310 BM337310
15	17	28.3	664	4	BM337310 BM337310
16	17	28.3	670	9	AG108977 AG108977
17	17	28.3	671	6	CA159481 CA159481
18	17	28.3	674	6	CA159393 CA159393
19	17	28.3	684	6	CT780149 CT780149
20	17	28.3	685	8	BM879942 BM879942
21	17	28.3	686	9	AG315012 AG315012
22	17	28.3	686	9	CT768499 CT768499
23	17	28.3	691	4	BM075656 BM075656
24	17	28.3	699	9	AG271710 AG271710

C 25	17	28.3	741	7	CF866719	CF866719	tr1c007xm
C 26	17	28.3	746	7	CF866367	CF866367	tr1c030xa
C 27	17	28.3	757	8	BZ775878	BZ775878	ih96609.b
C 28	17	28.3	794	6	CB896707	CB896707	tr1c007xm
C 29	17	28.3	800	6	CB902362	CB902362	tr1c030xa
C 30	17	28.3	805	6	CB677677	CB677677	OSJNCa10
C 31	17	28.3	820	9	CL805656	CL805656	OR_C8A001
C 32	17	28.3	834	9	CL858566	CL858566	OR_C8A008
C 33	17	28.3	841	8	BZ745534	BZ745534	OEEDC35TM
C 34	17	28.3	844	8	CG953853	CG953853	MBEDK10TF
C 35	17	28.3	848	8	BZ985406	BZ985406	PUGEV95TB
C 36	17	28.3	882	9	CL460518	CL460518	SA1L_106B
C 37	17	28.3	885	9	CG030301	CG030301	PUPOM18TD
C 38	17	28.3	938	9	CG217412	CG217412	OG3BP06TV
C 39	17	28.3	988	1	AV271892	AV271892	AV271892
C 40	17	28.3	991	9	CG234070	CG234070	OGVCP15TV
C 41	17	28.3	1792	9	CL514108	CL514108	SA1L_882
C 42	17	28.3	2864	3	AK048439	AK048439	Mus_muscu
C 43	16	26.7	172	7	CF746363	CF746363	UI-M-GVO-
C 44	16	26.7	211	1	AL915633	AL915633	AL915633
C 45	16	26.7	217	9	CR046864	CR046864	Reverse s

ALIGNMENTS

RESULT 1
LOCUS AV299623
DEFINITION AV299623 RIKEN full-length enriched, 8 days embryo Mus musculus
CDNA clone 5730470F16.3, mRNA sequence.
VERSION AV299623.1 GI:6332114
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirazane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Iwata, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koyama, S., Kusakabe, M., Matsuura, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugabe, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomihata, N., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

REFERENCE
AUTHORS
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirazane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Iwata, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koyama, S., Kusakabe, M., Matsuura, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugabe, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomihata, N., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
COMMENT

Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/
Saeki, N., Iwata, M., Watabiki, M., Ozawa, K., Tanaka, T., Itoneda, Y., Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh, M., Kiteunui, T., Akiyama, J., Shibata, K., Iwata, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.riken.go.jp) for

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 131.858 Seconds
(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-18

Perfect score: 59

Sequence: 1 CATTAACGAGTTATTCTTG.....TATTAATCACTAAGTCTT 59

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn19808:*
2: geneseqn19808:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20018:*
6: geneseqn20028:*
7: geneseqn20028:*
8: geneseqn20038:*
9: geneseqn20038:*
10: geneseqn20038:*
11: geneseqn20038:*
12: geneseqn20048:*
13: geneseqn20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	AAV23190	AAV23190 Lactococc
2	19	32.2	60	AAV23205	AAV23205 Lactococc
3	17	28.8	60	AAV23211	AAV23211 Lactococc
4	17	28.8	60	AAV23182	AAV23182 Lactococc
5	17	28.8	60	AAV23208	AAV23208 Lactococc
6	17	28.8	12	ADFS0636	ADFS0636 Murine DN
7	17	28.8	4	AA545384	AA545384 Chemical
8	17	28.8	4	AA546426	AA546426 Tumour su
9	17	28.8	6	ABK28226	ABK28226 DNA trans
10	17	28.8	13	ABD32966_02	ABD32966_02
11	16	27.1	59	AAV23210	AAV23210 Lactococc
12	16	27.1	60	AAV23177	AAV23177 Lactococc
13	16	27.1	60	AAV23191	AAV23191 Lactococc
14	16	27.1	60	AAV23197	AAV23197 Lactococc
15	16	27.1	60	AAV23189	AAV23189 Lactococc
16	16	27.1	60	AAV23200	AAV23200 Lactococc
17	16	27.1	60	AAV23214	AAV23214 Lactococc
18	16	27.1	60	AAV23198	AAV23198 Lactococc
19	16	27.1	60	AAV23180	AAV23180 Lactococc
20	16	27.1	60	AAV23199	AAV23199 Lactococc

21	16	27.1	265	2	AAQ23880	AAQ23880 PKTH1820
22	16	27.1	265	2	AAQ46205	AAQ46205 Promoter
23	16	27.1	265	2	AAT31875	AAT31875 Promoter
24	16	27.1	651	3	AAT13601	AAT13601 Aspergill
25	16	27.1	687	6	ABQ19920	ABQ19920 Oligonuc1
26	16	27.1	687	6	ABQ19921	ABQ19921 Oligonuc1
27	16	27.1	1236	10	ADH84590	ADH84590 Enterococ
28	16	27.1	2000	8	ADA72746	ADA72746 Rice gene
29	16	27.1	2286	4	AAI37033	AAI37033 Human mus
30	16	27.1	2286	4	ABK60021	ABK60021 cDNA enco
31	16	27.1	2286	12	ADJ30771	ADJ30771 Human mus
32	16	27.1	3660	4	AB110174	AB110174 Drosophi1
33	16	27.1	4674	4	AAI37031	AAI37031 Human mus
34	16	27.1	4674	8	ABK60019	ABK60019 cDNA enco
35	16	27.1	4674	12	ADJ30769	ADJ30769 Human mus
36	16	27.1	5300	6	ABK39947	ABK39947 Human che
37	16	27.1	5300	6	AB132549	AB132549 Human imm
38	16	27.1	28564	10	ADRS9178	ADRS9178 Human gen
39	16	27.1	32768	2	AAI33060	AAI33060 Enterococ
40	16	27.1	32768	6	ABK98855	ABK98855 Enterococ
41	16	27.1	110000	6	ABK90521_02	ABK90521_02
42	16	27.1	110000	10	ADP77343_15	ADP77343_15
43	16	27.1	161280	10	ADD67054	ADD67054 Human lun
44	16	27.1	161280	10	ADK88308	ADK88308 Human lun
45	16	27.1	179487	12	ADQ97687	ADQ97687 Human can

ALIGNMENTS

RESULT 1

ID AAV23190 standard; DNA; 59 BP.

AAV23190;

28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cp21.

Lactococcus lactis; constitutional promoter; optimise; spacer;

artificial promoter library; gene expression; de.

Synthetic.

Lactococcus lactis.

Key

promoter

FT

PN

PD

PF

PR

PA

XX

XX

XX

XX

XX

Location/Qualifiers

4..59

/tag= a

/standard_name= "Constitutional promoter"

WO9807846-A1.

26-FEB-1998.

25-AUG-1997;

97WO-DK000342.

23-AUG-1996;

96DK-00000886.

(JENS/) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 49; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 458.474 Seconds
(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657b-18

Sequence: 1 CATTACCGAGTTTATCTTCTG.....TATTAATCTATAGTACTGTT 59

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 .segs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_pi:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	BD006949	BD006949 Artificia
2	20	33.9	148481	CR548627	CR548627 Dario rer
3	19	32.2	60	BD006964	BD006964 Artificia
4	19	32.2	130764	AC146932	AC146932 Sus scrofa
5	19	32.2	189228	AC151176	AC151176 Bos tauri
6	18	30.5	573	AYE252319	AYE252319 Agriodril
7	18	30.5	31988	AF040642	AF040642 Caenorhab
8	18	30.5	147571	AC080177	AC080177 Homo sapi
9	18	30.5	169128	AP005864	AP005864 Oryza sat
10	18	30.5	198959	AL354806	AL354806 Human DNA
11	18	30.5	210088	AC116952	AC116952 Mus muscu
12	18	30.5	218092	AC022238	AC022238 Homo sapi
13	18	30.5	251834	CR847858	CR847858 Dario rer
14	18	30.5	274498	AC006765	AC006765 Caenorhab
15	17	28.8	60	BD006941	BD006941 Artificia
16	17	28.8	60	BD006967	BD006967 Artificia
17	17	28.8	60	BD006970	BD006970 Artificia
18	17	28.8	5689	AX251180	AX251180 Sequence
19	17	28.8	5689	AX277927	AX277927 Sequence

20	17	28.8	5689	6	AX323612	AX323612 Sequence
21	17	28.8	5867	2	AC020180	AC020180 Drosophila
22	17	28.8	7731	14	HP085660	HP085660 Human papill
23	17	28.8	10753	1	AE011132	AE011132 Methanosa
24	17	28.8	19789	5	AY728221	AY728221 Batrachos
25	17	28.8	28324	8	SPC1126	AL034490 S.pombe c
26	17	28.8	33417	9	AC104537	AC104537 Homo sapi
27	17	28.8	40662	3	CEY328128	AL031632 Caenorhab
28	17	28.8	44219	8	AC005512	AC005512 Homo sapi
29	17	28.8	96561	8	AP006709	AP006709 Lotus cor
30	17	28.8	110885	8	AP006701	AP006701 Lotus cor
31	17	28.8	114252	5	BX649441	BX649441 Zebrafish
32	17	28.8	128242	9	AC112172	AC112172 Homo sapi
33	17	28.8	130632	9	AC004547	AC004547 Homo sapi
34	17	28.8	140384	9	AC106784	AC106784 Homo sapi
35	17	28.8	142227	2	AL162492	AL162492 Homo sapi
36	17	28.8	143841	9	AC008964	AC008964 Homo sapi
37	17	28.8	148279	9	AC012472	AC012472 Homo sapi
38	17	28.8	162785	5	BX547938	BX547938 Zebrafish
39	17	28.8	166155	2	AC093129	AC093129 Papio anu
40	17	28.8	167573	10	AC121587	AC121587 Mus muscu
41	17	28.8	168562	10	AC147262	AC147262 Mus muscu
42	17	28.8	173742	2	AC068626	AC068626 Homo sapi
43	17	28.8	174999	5	AL840636	AL840636 Zebrafish
44	17	28.8	175343	2	AL807799	AL807799 Dario rer
45	17	28.8	176195	3	AC012165	AC012165 Drosophila

ALIGNMENTS

RESULT 1
BD006949
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD006949
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006949.1 GI:18635320
JP 2001503249-A/18.
Lactococcus lactis
Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PI

PC

CC

FT

FEATURES

ORIGIN

Query Match 100.0%; Score 59; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 6.2e-24;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATTACCGAGTTTATCTTGACACCGTTATCGGGTGTATTAATAGTACTGTT 59
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1206.06 Seconds
(Without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-18

Perfect score: 59
Sequence: 1 CATACCGAGTTATCTCTG.....TATATACCTACTACTGTT 59

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 segs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	21	35.6	653 8	CC181212 JSa1BAC 0
C 2	21	35.6	691 8	CC385473 PUM1W26TB
C 3	21	35.6	934 8	CC43048 PUMR15TD
C 4	18	30.5	371 2	AM432589 sh7604.y
C 5	18	30.5	392 2	BE659450 GM700009B
C 6	18	30.5	453 2	BF113638 EST441228
C 7	18	30.5	675 2	BH590409 BOCY142TR
C 8	18	30.5	691 5	BH544356 GM860002B
C 9	18	30.5	717 8	BH956530 Od35110.
C 10	18	30.5	724 8	BH590296 BOC226TF
C 11	18	30.5	844 9	CG953853 MBEDK10TF
C 12	17	28.8	150 1	AA249342 J0381.1 Af
C 13	17	28.8	249 9	CM64743 8503.1 Af
C 14	17	28.8	282 9	CG670268 OS1469470
C 15	17	28.8	345 6	BH713076 BOHV52TR
C 16	17	28.8	368 6	CB501545 ssal1ra020
C 17	17	28.8	419 1	AA506935 nh64a1.8
C 18	17	28.8	420 4	BI010811 MR2-EN009
C 19	17	28.8	454 1	AA493266 nh55b03.8
C 20	17	28.8	508 7	CV503300 69754.1 M
C 21	17	28.8	539 8	AO971139 RPT-25-3
C 22	17	28.8	592 5	BX312577 BX312577
C 23	17	28.8	628 9	CB335717 C19F-988-
C 24	17	28.8	632 9	CL596215 OB_Ba005

C 25	17	28.8	643 6	CA051009 ssal1rga51
C 26	17	28.8	658 6	CA347397 678510 NC
C 27	17	28.8	674 8	BH794142 ME MBa000
C 28	17	28.8	689 8	BZ082441 11e41b06.
C 29	17	28.8	694 9	AG342781 Mus muscu
C 30	17	28.8	732 8	AZ185046 SF_1004_A
C 31	17	28.8	732 8	AG604469 Mus muscu
C 32	17	28.8	755 5	BX113930 BX113930
C 33	17	28.8	761 9	CR301077 Medicago
C 34	17	28.8	846 9	CG893639 03S041-0
C 35	17	28.8	822 8	CC105362 CSU-R34.1
C 36	17	28.8	1027 4	BG622307 602646974
C 37	17	28.8	1658 4	BG327215 602426161
C 38	17	28.8	2864 3	AK048439 Mus muscu
C 39	16	27.1	240 1	AV253915 AV253915
C 40	16	27.1	241 5	BX726776 BX726776
C 41	16	27.1	243 1	AV234563 AV234563
C 42	16	27.1	243 1	AV275959 AV275959
C 43	16	27.1	258 2	BB307304 BB307304
C 44	16	27.1	258 4	BI276919 UI-R-CW0-
C 45	16	27.1	259 6	CD947316 RFB_87 Ge

ALIGNMENTS

RESULT 1
LOCUS CC181212/c 653 bp DNA linear GSS 05-MAY-2003
DEFINITION JSa1BAC_0 G19 T7 Sal I (10-15 KB) BAC library Zea mays genomic,
genomic survey sequence.
ACCESSION CC181212
VERSION CC181212.1 GI:30353258
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogonaceae; Zea.
1 (bases 1 to 653)
Yuan,Y., Samuigne,P.J. and Bennetzen,J.L.
Methylation-spanning linker libraries link gene-rich regions and
identify epigenetic boundaries in Zea mays
Genome Res. 12 (9), 1345-1349 (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Bennetzen JL
Department of Genetics
University of Georgia
1057 Green Street, Athens, GA 30602, USA
Tel: 706 542 9729
Fax: 706 583 0972
Email: maize@uga.edu
Seq primer: T7
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..653
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/dev_stage="10 days seedling"
/lab_host="DH10B"
/clone_lib="Sal I (10-15 KB) BAC library"
/note="Organ: Young leaves; Vector: pBlacR11"

ORIGIN

Query Match 35.6%; Score 21; DB 8; Length 653;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

31 TCGGGGTTGTAATACTACTA 51
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds

(without alignments)
2646.796 Million cell updates/sec

Title: US-09-242-657b-19

Perfect score: 60
Sequence: 1 CATGTAGAGAGTTATCTTG.....TATAATCTCAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23191
2	20	33.3	60	2	AAV23177
3	20	33.3	755	6	ABN99221
4	20	33.3	1062	6	ABZ13652
5	20	33.3	1062	6	ADG87654
6	20	33.3	1062	6	ADG87655
7	20	33.3	1062	6	ADA68053
8	20	33.3	1305	3	AAC47953
9	20	33.3	1305	3	AAC51562
10	20	33.3	2118	3	AAC51464
11	18	30.0	447	6	ABN94813
12	18	30.0	1055	6	ABL89967
13	18	30.0	1203	8	ACC62477
14	18	30.0	1215	8	ACC62497
15	18	30.0	1381	11	ADM2882
16	18	30.0	1390	10	ADB63044
17	18	30.0	1876	8	ABX05008
18	18	30.0	1936	8	ADB63040
19	18	30.0	1962	12	ADQ24783
20	18	30.0	2096	12	ADQ64791

21	17	28.3	59	2	AAV23210	AAV23210 Lactococc
22	17	28.3	60	2	AAV23211	AAV23211 Lactococc
23	17	28.3	60	2	AAV23182	AAV23182 Lactococc
24	17	28.3	60	2	AAV23189	AAV23189 Lactococc
25	17	28.3	60	2	AAV23200	AAV23200 Lactococc
26	17	28.3	60	2	AAV23198	AAV23198 Lactococc
27	17	28.3	60	2	AAV23199	AAV23199 Lactococc
28	17	28.3	296	3	AAC03364	AAV23199 Lactococc
29	17	28.3	421	5	AAV68679	AAV23199 Lactococc
30	17	28.3	421	6	ABK38590	AAV23199 Lactococc
31	17	28.3	421	8	ACH10919	AAV23199 Lactococc
32	17	28.3	421	8	ABX99870	AAV23199 Lactococc
33	17	28.3	421	10	ADH46133	AAV23199 Lactococc
34	17	28.3	421	12	ADH71885	AAV23199 Lactococc
35	17	28.3	421	13	ADJ20052	AAV23199 Lactococc
36	17	28.3	805	4	AAI97318	AAV23199 Lactococc
37	17	28.3	1199	10	ADJ22622	AAV23199 Lactococc
38	17	28.3	2094	8	ACA27846	AAV23199 Lactococc
39	17	28.3	6294	2	AAV31191	AAV23199 Lactococc
40	17	28.3	11165	3	AAI5186	AAV23199 Lactococc
41	16	26.7	58	2	AAV23184	AAV23184 Lactococc
42	16	26.7	58	6	ABK98625	AAV23184 Lactococc
43	16	26.7	58	9	ACD13876	AAV23184 Lactococc
44	16	26.7	59	2	AAV23190	AAV23190 Lactococc
45	16	26.7	59	2	AAV23185	AAV23185 Lactococc

ALIGNMENTS

RESULT 1
AAV23191 standard; DNA; 60 bp.

AAV23191;

28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cp23.

Lactococcus lactis; constitutional promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Synthetic.

Lactococcus lactis.

Key Location/Qualifiers
promoter 4..60
/*tag= a
/standard_name= "Constitutional promoter"

W09807846-A1.

26-FEB-1998.

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENS/) JENSEN P R.

Jensen PR, Hammer K;

WPI; 1998-179062/16.

New artificial promoter libraries - containing consensus promoter sequences and variable spacers, used to generate promoters for optimising expression of genes.

Claim 28; Page 49; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds
(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-19

Perfect score: 60

Sequence: 1 CATGTAGAGATTATCTTCTTG.....TATTAATCTCAGTACTGTT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenBank:

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006950	BD006950 Artificial
2	20	33.3	60	BD006936	BD006936 Artificial
3	20	33.3	715	AF083684	AF083684 Arabidops
4	20	33.3	1062	AX412332	AX412332 Sequence
5	20	33.3	1062	AX412333	AX412333 Sequence
6	20	33.3	1062	AX506762	AX506762 Sequence
7	20	33.3	1062	AX651506	AX651506 Sequence
8	20	33.3	1062	AY079106	AY079106 Arabidops
9	20	33.3	1110	AF339700	AF339700 Arabidops
10	20	33.3	1189	AF324700	AF324700 Arabidops
11	20	33.3	1206	AF326880	AF326880 Arabidops
12	20	33.3	1210	AF419569	AF419569 Arabidops
13	20	33.3	1305	AY087926	AY087926 Arabidops
14	20	33.3	85109	ATF2U13	ATF2U13 Arabidops
15	20	33.3	129757	ATF2K15	ATF2K15 Arabidops
16	19	31.7	1236	BV171889	BV171889 Bgmu6792
17	19	31.7	1236	ATPRXCB	ATPRXCB Arabidops
18	19	31.7	226554	AC108956	AC108956 Rattus no
19	19	31.7	238316	AC094974	AC094974 Rattus no

20	18	30.0	164	11	G20291	G20291 human STS A
21	18	30.0	191	11	BV079516	BV079516 human STS B
22	18	30.0	321	11	G25574	G25574 human STS E
23	18	30.0	447	6	AX408664	AX408664 Sequence
24	18	30.0	455	11	G29541	G29541 human STS S
25	18	30.0	1381	6	AX834443	AX834443 Sequence
26	18	30.0	1381	9	AK096905	AK096905 Homo sapi
27	18	30.0	1390	6	AX747673	AX747673 Sequence
28	18	30.0	1390	9	AK092692	AK092692 Homo sapi
29	18	30.0	1446	9	BC017926	BC017926 Homo sapi
30	18	30.0	1936	6	AX747669	AX747669 Sequence
31	18	30.0	1936	9	AK092682	AK092682 Homo sapi
32	18	30.0	2096	6	CO843305	CO843305 Sequence
33	18	30.0	2096	9	AX124273	AX124273 Homo sapi
34	18	30.0	27049	9	AL138842	AL138842 Homo sapi
35	18	30.0	72170	9	AL445068	AL445068 Human DNA
36	18	30.0	82895	2	OSIG00041	OSIG00041 Human DNA
37	18	30.0	93003	9	AF547386	AF547386 Homo sapi
38	18	30.0	103274	2	CR391972	CR391972 Homo sapi
39	18	30.0	107315	9	AL136303	AL136303 Human DNA
40	18	30.0	112897	2	OSIG00042	OSIG00042 Human DNA
41	18	30.0	120896	9	AL161790	AL161790 Human DNA
42	18	30.0	133413	2	EX548022	EX548022 Dantio rer
43	18	30.0	143754	2	AL359093	AL359093 Homo sapi
44	18	30.0	144670	8	OSUN00084	OSUN00084 Oryza sat
45	18	30.0	148982	2	CR361556	CR361556 Dantio rer

ALIGNMENTS

RESULT 1
LOCUS BD006950
DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION BD006950
VERSION BD006950.1 GI:18635321
KEYWORDS JP 2001503249-A/19.
SOURCE Lactococcus lactis
ORGANISM Lactococcus lactis
REFERENCE 1 (bases 1 to 60)
AUTHORS Hammer, K. and Janssen, P.R.
TITLE Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL Patent: JP 2001503249-A 19 13-MAR-2001,
PETER RUDAR JANSSEN
COMMENT OS Lactococcus lactis
PN JP 2001503249-A/19
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key promoter location/Qualifiers
FT 1..60 location/Qualifiers
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0% Score 60; DB 6; Length 60;
Best Local Similarity 100.0% Pred. No. 1.4e-22;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGTAGAGATTATCTTCTTGACGATTAGTAGGGGCTGTATTAATCTCAGTACTGTT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657b-19

Sequence: 1 CATGTAGAGATTATCTCTG.....TATATATCTCAGTACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	33.3	421	1	AV549300
2	20	33.3	476	1	AV544323
3	20	33.3	477	1	AV544323
4	20	33.3	477	1	AV544323
5	20	33.3	477	1	AV544323
6	20	33.3	477	1	AV544323
7	20	33.3	477	1	AV544323
8	20	33.3	477	1	AV544323
9	20	33.3	477	1	AV544323
10	20	33.3	477	1	AV544323
11	20	33.3	477	1	AV544323
12	20	33.3	477	1	AV544323
13	20	33.3	477	1	AV544323
14	20	33.3	477	1	AV544323
15	20	33.3	477	1	AV544323
16	20	33.3	477	1	AV544323
17	20	33.3	477	1	AV544323
18	20	33.3	477	1	AV544323
19	20	33.3	477	1	AV544323
20	20	33.3	477	1	AV544323
21	20	33.3	477	1	AV544323
22	20	33.3	477	1	AV544323
23	20	33.3	477	1	AV544323
24	20	33.3	477	1	AV544323

C	25	20	33.3	620	1	AV441260	AV441260
	26	20	33.3	621	6	CB255682	CB255682
	27	20	33.3	631	6	CB260269	CB260269
	28	20	33.3	634	6	CB260872	CB260872
	29	20	33.3	647	6	CB260750	CB260750
	30	20	33.3	734	6	CB261136	CB261136
	31	20	33.3	802	7	CK119274	CK119274
	32	19	31.7	637	9	CR829063	CR829063
	33	19	31.7	722	9	CB839399	CB839399
	34	19	31.7	781	8	BH532824	BH532824
	35	18	30.0	189	7	F00348	F00348
	36	18	30.0	196	1	AT243087	AT243087
	37	18	30.0	232	7	F35827	F35827
	38	18	30.0	235	1	AA181515	AA181515
	39	18	30.0	236	7	F00368	F00368
	40	18	30.0	258	7	F26701	F26701
	41	18	30.0	268	7	Z28893	Z28893
	42	18	30.0	358	7	F34953	F34953
	43	18	30.0	409	1	AI084581	AI084581
	44	18	30.0	415	1	AI580838	AI580838
	45	18	30.0	422	1	AA936591	AA936591

ALIGNMENTS

RESULT 1
LOCUS AV549300 421 bp mRNA linear EST 20-FEB-2004
DEFINITION AV549300 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone R206a02R 5', mRNA sequence.
AV549300
AV549300.1 GI:8720713

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 421)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Asanizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)

CONTACT: Erika Asanizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source
location/Qualifiers
1..421
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="R206a02R"
/issue_type="Roots"
/clone_id="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 33.3% Score 20; DB 1; Length 421;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 9 AGTTATCTTGCAGATTA 28
DB 330 AGTTATCTTGCAGATTA 349

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 134.093 Seconds

(without alignments)
2648.796 Million cell updates/sec

Title: US-09-242-657B-20

Perfect score: 60
Sequence: 1 CATGGGTAAGTTATCTTC.....TATATAGAGTACTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23192 Lactococc
2	18	30.0	62	5	ABV51941 Human pro
3	17	28.3	59	2	AAV23202 Lactococc
4	17	28.3	60	2	AAV23201 Lactococc
5	17	28.3	32	8	ABZ17722 S2 subtra
6	17	28.3	958	6	ABZ14910 Arabidops
7	17	28.3	1596	4	ABA08852 Human sec
8	17	28.3	2780	13	ADR25834 Contamination (11 o
9	17	28.3	89736	13	ABD32956_10 Contamination (11 o
10	16	26.7	328	12	ADP62471 Maize car
11	16	26.7	477	4	ABA58405 Human foe
12	16	26.7	477	4	AAI38045 Probe #67
13	16	26.7	477	4	AAK32194 Human bon
14	16	26.7	477	4	ABG31893 Human liv
15	16	26.7	477	6	ABS06962 Human gen
16	16	26.7	1555	2	AAI6243 Arabidops
17	16	26.7	208765	13	ADQ97430 Mouse can
18	16	26.7	301477	12	ADQ97362 Human can
19	15	25.0	131	7	ADG68868 Corn seed
20	15	25.0	210	12	ADK16498 Nanarcha

21	15	25.0	231	8	ACA38340 Prokaryot
22	15	25.0	234	8	ACA40378 Prokaryot
23	15	25.0	320	3	AAC31553 Human sec
24	15	25.0	332	2	AAZ33865 Tobacco p
25	15	25.0	417	3	AAC24643 Human sec
26	15	25.0	446	3	AAC05175 Human sec
27	15	25.0	461	9	ACH38741 Human foe
28	15	25.0	472	9	ACH27865 Human adu
29	15	25.0	492	9	ACH36651 Human end
30	15	25.0	510	9	ACH37898 Human end
31	15	25.0	549	1	AAH13608 Human CDN
32	15	25.0	568	4	AAH76698 Human sec
33	15	25.0	573	3	AAZ65311 Human sec
34	15	25.0	573	10	ABE11719 Human sec
35	15	25.0	655	13	ADQ49196 Novel can
36	15	25.0	808	2	AAQ26089 Chicken e
37	15	25.0	823	5	AAQ33986 Human CDN
38	15	25.0	823	5	ABN74743 Bovine em
39	15	25.0	918	6	ABN74743 Bovine em
40	15	25.0	954	6	ABN74743 Bovine em
41	15	25.0	954	6	ABN74743 Bovine em
42	15	25.0	1015	6	ABK51443 cDNA enco
43	15	25.0	1025	12	ADQ97366 Human can
44	15	25.0	1112	10	ADC20846 Human sec
45	15	25.0	1112	10	ABT16983 Human sec

ALIGNMENTS

RESULT 1
AAV23192
ID AAV23192 standard; DNA, 60 BP.

AAV23192;
28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cp24.

Lactococcus lactis; constitutional promoter; optimise; spacer;
artificial promoter library; gene expression; ds.

Synthetic.
Lactococcus lactis.

Key Location/Qualifiers
promoter 4..60
/*tag= a
/standard_name= "Constitutional promoter"

W09807846-A1.

26-FEB-1998.

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENS/) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter
sequences and variable spacers, used to generate promoters for optimizing
expression of genes.

Claim 28; Page 50; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the
construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 05:21:56 ; Search time 466.245 Seconds

(without alignments)
6235.594 Million cell updates/sec

Title: US-09-242-657B-20

Sequence: 1 CATGGGTAAGTTATTCTTC.....TATAATAAGTACTACTGTT 60

Scoring table: OLIGO_NUC

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

```

1:  gb_da:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_ey:*
13: gb_un:*
14: gb_vi:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	6	BD006951	BD006951 Artificia
2	20	33.3	67845	6	AC083980	AC083980 Homo sapi
3	20	33.3	121029	9	AL391726	AL391726 Human DNA
c 4	20	33.3	142870	2	AC011888	AC011888 Homo sapi
c 5	19	31.7	173382	10	AC124352	AC124352 Mus musc
6	19	31.7	174678	2	AC134449	AC134449 Mus musc
7	19	31.7	245070	2	AC129467	AC129467 Rattus no
c 8	18	30.0	54982	6	C0520093	C0520093 Sequence
9	18	30.0	106123	9	AC008910	AC008910 Homo sapi
c 10	18	30.0	106123	9	AL158015	AL158015 Human DNA
11	18	30.0	108799	9	AC090989	AC090989 Homo sapi
12	18	30.0	118150	9	HSJ800CC24	HSJ800CC24 Human DNA
c 13	18	30.0	159524	9	CNS01D00	AL115193 Human chr
c 14	18	30.0	173834	9	AC108044	AC108044 Homo sapi
c 15	18	30.0	176148	9	AC010798	AC010798 Homo sapi
c 16	18	30.0	176825	2	AP001203	AP001203 Homo sapi
c 17	18	30.0	183019	2	AC024388	AC024388 Homo sapi
18	18	30.0	242425	2	AC136818	AC136818 Rattus no
19	17	28.3	59	6	BD006961	BD006961 Artificia

20	17	28.3	60	6	BD006960	BD006960
21	17	28.3	800	8	BT000376	BT000376 Arabidops
22	17	28.3	958	6	AK508020	AK508020 Sequence
23	17	28.3	1043	3	AF438024	AF438024 Asclepias
24	17	28.3	1266	8	AY919139	AY919139 Heteroig
25	17	28.3	1801	8	AF370486	AF370486 Arabidops
26	17	28.3	2780	9	AK000163	AK000163 Homo sapi
27	17	28.3	17144	1	AE0001473	AE0001473 Helicobac
28	17	28.3	23606	9	HSU71B4	Z69734 Human DNA s
29	17	28.3	31151	2	AC107201_8	Continuation (9 of
30	17	28.3	37013	3	CER58H1	Z75954 Ctenorhadi
31	17	28.3	51531	2	AC115089	AC115089 Homo sapi
32	17	28.3	62524	2	AC091116	AC091116 Homo sapi
33	17	28.3	65954	2	AC069578	AC069578 Homo sapi
34	17	28.3	81188	2	AP001128	AP001128 Homo sapi
35	17	28.3	87941	10	AL954374	AL954374 Mouse DNA
36	17	28.3	88011	8	ATT16L4	ALU79344 Arabidops
37	17	28.3	94870	9	AL157948	AL157948 Human DNA
38	17	28.3	105168	8	AC091317	AC091317 Homo sapi
39	17	28.3	106312	2	AC110832_4	Continuation (5 of
40	17	28.3	110000	2	AC107201_7	Continuation (8 of
41	17	28.3	110000	2	AC110828_1	Continuation (2 of
42	17	28.3	110000	2	AC112373_01	Continuation (2 of
43	17	28.3	116552	9	AC004022	AC004022 Homo sapi
44	17	28.3	118103	9	AC009817	AC009817 Homo sapi
45	17	28.3	127061	2	AC118553	AC118553 Mus muscu

ALIGNMENTS

RESULT 1
BD006951

DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.

VERSION BD006951.1 GI:18635322
KEYWORDS JP 2001503249-A/20.

ORGANISM

REFERENCE

TITLE

JOURNAL Patent: JP 2001503249-A 20 13-MAR-2001;

COMMENT

FEATURES

ORIGIN

Query

Match

Qy

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1226.5 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-20

Perfect score: 60
Sequence: 1 CATGGGTAAGTATTCTCTC.....TATATATAGACTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	649	9	CR189011 Reverse 8
2	19	31.7	684	9	CR232325 Reverse 8
3	19	31.7	686	9	BX963788 Reverse 8
4	18	30.0	710	8	BH952660 cdl68e07.
5	18	30.0	737	7	CK308567 SB02046B2
6	18	30.0	744	8	BZ040626 lka47c02.
7	18	30.0	776	9	CL663719 PR10145a.
8	17	28.3	254	6	C04909 C04909 Huma
9	17	28.3	270	4	BM662008 UI-E-CX1-
10	17	28.3	318	1	AI242361 q115d06.x
11	17	28.3	326	9	CL341881 RPCI44.26
12	17	28.3	333	6	CD592886 RK0744AG0
13	17	28.3	356	1	AA776566 ah14g06.8
14	17	28.3	359	4	AI973211 wt53b02.x
15	17	28.3	359	1	BM692518 UI-E-CU1-
16	17	28.3	363	5	BU733024 UI-E-CU1-
17	17	28.3	400	1	AV619627 AV819627
18	17	28.3	404	5	BA350193 ht11a03.x
19	17	28.3	404	5	BY491519 BY491519
20	17	28.3	411	5	BQ125088 pb31e08.Y
21	17	28.3	422	8	BH382695 AG-ND-167
22	17	28.3	427	7	W94555 ze04c07.a1
23	17	28.3	438	4	BI511556 BBI60006A
24	17	28.3	439	5	BP585563 BP585563

25	17	28.3	443	8	AQ017207 CTT-HSP-2
26	17	28.3	453	2	BE932882 CM2-HT087
27	17	28.3	478	1	AV559015 AV559015
28	17	28.3	521	9	AZ161180 SP 0069 A
29	17	28.3	524	9	CL003689 gbl1 VMS
30	17	28.3	525	5	BK565243 BK565243
31	17	28.3	532	6	CB471902 sn46_F12.
32	17	28.3	581	2	AM177217 CM2-CT012
33	17	28.3	587	8	BZ938235 CH240.63A
34	17	28.3	591	6	CD512881 AGENCOURT
35	17	28.3	592	5	BU729331 UI-E-CU1-
36	17	28.3	605	4	BM670063 UI-E-DX1-
37	17	28.3	629	8	AQ940350 Sheared D
38	17	28.3	637	9	DR15G23S Danio rer
39	17	28.3	638	6	CB851465 UI-CF-DU1
40	17	28.3	639	5	BUC08057 UI-CF-FNO
41	17	28.3	651	5	BM978949 UI-CF-DU1
42	17	28.3	652	8	BZ308825 i475e08.b
43	17	28.3	652	9	CE751910 L19T-988-
44	17	28.3	657	4	BG185199 RST4137 A
45	17	28.3	657	6	CB306345 UI-CF-EN1

ALIGNMENTS

RESULT 1	CR189011	649 bp	DNA	linear	GSS 06-JUL-2004
DEFINITION	Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN349c07, genomic survey sequence.				
ACCESSION	CR189011	GI:49967860			
VERSION	GSS; genome survey sequence; M1CER.				
KEYWORDS	GSS; Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 649)				
AUTHORS	Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J., and Bradley,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/M1CER				
FEATURES	Location/Qualifiers				
source	1..649				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	/clone="MHPN349c07"				
	/clone_1db="MHPN"				
ORIGIN					
Query Match	31.7%; Score 19; DB 9; Length 649;				
Best Local Similarity	100.0%; Pred. No. 4.5;				
Matches	19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	5 GGTAGTTATTCTTCA 23				
DB	254 GGTAGTTATTCTTCA 272				
RESULT 2	CR232325	684 bp	DNA	linear	GSS 06-JUL-2004
DEFINITION	Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN75n04, genomic survey sequence.				
ACCESSION	CR232325	GI:50011174			
KEYWORDS	GSS; genome survey sequence; M1CER.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 04:31:36 ; Search time 131.858 Seconds
(without alignments)

2648.796 Million cell updates/sec

Title: US-09-242-657B-21

Perfect score: 59
Sequence: 1 CTTTGACAGTTTATCTCTGA.....TATATACACATGACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn19808:*\n2: geneseqn19908:*\n3: geneseqn20008:*\n4: geneseqn20018:*\n5: geneseqn20018:*\n6: geneseqn20028:*\n7: geneseqn20028:*\n8: geneseqn20038:*\n9: geneseqn20038:*\n10: geneseqn20038:*\n11: geneseqn20048:*\n12: geneseqn20048:*\n13: geneseqn20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	AAV23193	AAV23193 Lactococc
2	59	100.0	12739	ABK98592	ABK98592 Vector pE
3	59	100.0	12739	ACD13843	ACD13843 Plasmid p
4	52	88.1	10929	ABK98591	ABK98591 pBP25 vec
5	52	88.1	10929	ACD13842	ACD13842 Plasmid p
6	35	59.3	60	ABK98622	ABK98622 Lambda CP
7	35	59.3	60	ACD13873	ACD13873 L. lactis
8	35	59.3	95	ABK98612	ABK98612 Lambda CP
9	35	59.3	95	ACD13863	ACD13863 L. lactis
10	35	59.3	9437	ABK98601	ABK98601 Vector pE
11	35	59.3	9437	ACD13852	ACD13852 Xyl opera
12	18	30.5	58	AAV23186	AAV23186 Lactococc
13	17	28.8	59	AAV23206	AAV23206 Lactococc
14	17	28.8	60	AAV23194	AAV23194 Lactococc
15	17	28.8	60	AAV23188	AAV23188 Lactococc
16	17	28.8	60	AAV23212	AAV23212 Lactococc
17	17	28.8	169659	12	ADQ59434 Human can
18	16	27.1	58	ABK98625	ABK98625 L. lactis
19	16	27.1	58	ACD13876	ACD13876 L. lactis
20	16	27.1	59	AAV23210	AAV23210 Lactococc

21	16	27.1	60	2	AAV23200	AAV23200 Lactococc
22	16	27.1	60	2	AAV23199	AAV23199 Lactococc
23	16	27.1	64	6	ABK98604	ABK98604 L. lactis
24	16	27.1	64	6	ACD13855	ACD13855 L. lactis
25	16	27.1	65	6	ABK98605	ABK98605 L. lactis
26	16	27.1	65	9	ACD13856	ACD13856 L. lactis
27	16	27.1	93	6	ABK98615	ABK98615 L. lactis
28	16	27.1	93	6	ACD13866	ACD13866 L. lactis
29	16	27.1	115	6	ABK98626	ABK98626 L. lactis
30	16	27.1	115	9	ACD13877	ACD13877 L. lactis
31	16	27.1	150	6	ABK98616	ABK98616 L. lactis
32	16	27.1	150	9	ACD13867	ACD13867 L. lactis
33	16	27.1	167	3	AACT78401	AACT78401 Human nar
34	16	27.1	443	5	ABK11399	ABK11399 Human nar
35	16	27.1	538	6	ABK62126	ABK62126 Human foe
36	16	27.1	538	4	AA142073	AA142073 Probe #10
37	16	27.1	538	4	AAK36347	AAK36347 Human bon
38	16	27.1	538	4	AAK10445	AAK10445 Human bra
39	16	27.1	538	4	ABK36027	ABK36027 Human liv
40	16	27.1	538	6	ABK10407	ABK10407 Human gen
41	16	27.1	557	6	ABK60231	ABK60231 Human col
42	16	27.1	755	6	ABK99221	ABK99221 Arabidops
43	16	27.1	1062	6	ABK13652	ABK13652 Arabidops
44	16	27.1	1062	6	ADG87654	ADG87654 A. thalia
45	16	27.1	1062	6	ADG87655	ADG87655 A. thalia

ALIGNMENTS

RESULT 1	AAV23193	AAV23193 standard; DNA; 59 BP.
ID	AAV23193	AAV23193
XX	AAV23193;	
AC	28-JUL-1998	(first entry)
DT	Lactococcus lactis	constitutional promoter Cp25.
XX	Lactococcus lactis;	constitutional promoter; optimise; spacer;
XX	artificial promoter library;	gene expression; ds.
KW	Synthetic.	
XX	Lactococcus lactis.	
OS	Lactococcus lactis.	
XX	Key	Location/Qualifiers
FT	promoter	3..59
FT	promoter	/*tag= a
FT	promoter	/standard_name= "Constitutional promoter"
PN	MO9807846-A1.	
XX	26-FEB-1998.	
PD	25-AUG-1997;	97WO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
PR	(JENSEN/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
PI	WPI; 1998-179062/16.	
DR	New artificial promoter libraries -	containing consensus promoter
PT	sequences and variable spacers,	used to generate promoters for optimising
PT	expression of genes.	
XX	Claim 28; Page 50; 89pp; English.	
PS	This is a Lactococcus lactis	constitutional promoter sequence used in the
XX	construction of an artificial promoter	library of the invention. The
CC		

1 CTTGGCAGTTTATCTTGACATGTAAGAGGGGCTGGTATATCACATAGTACTGTT 55

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 06:09:45 ; Search time 1206.06 Seconds
(without alignments)
1862.095 Million cell updates/sec

Title: US-09-242-657B-21

Perfect score: 59

Sequence: 1 CTTGGCAGTTATCTTGA.....TATTAATCAGTACTGTT 59

Scoring table: OLIGO_NTC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	18	30.5	445	9	CL361142 RPTC14_34
2	18	30.5	490	8	BH912345 PST1271-2
3	18	30.5	570	9	CL385164 RPTC14_33
4	18	30.5	622	6	CB508948 sealc028
5	18	30.5	718	5	BX910735 BX910735
6	18	30.5	899	9	CNS030PT
7	18	30.5	1042	5	B066558 Tetradon
8	18	30.5	1733	9	CL022502
9	17	28.8	395	8	BH356393 CH230-82D
10	17	28.8	401	2	BF758229 CM3-CT057
11	17	28.8	434	4	BJS98819
12	17	28.8	437	7	N58914
13	17	28.8	582	2	AV967251
14	17	28.8	584	7	CO481889
15	17	28.8	601	4	B0313250
16	17	28.8	615	4	BJS90096
17	17	28.8	626	7	CO207763
18	17	28.8	627	4	BJ602198
19	17	28.8	634	5	BW274089
20	17	28.8	640	8	AZ378371
21	17	28.8	641	9	CL003396
22	17	28.8	656	8	BZ38053
23	17	28.8	708	4	BJ165556
24	17	28.8	719	4	BJS84408

25	17	28.8	719	9	AG218986	AG218986 Drosophila
26	17	28.8	738	4	BJ604414	BJ604414 B604414
27	17	28.8	763	4	BU611100	BU611100 BU611100
28	17	28.8	772	8	CC409804	CC409804 PPHIP59TD
29	17	28.8	789	9	BH694879	BH694879 BOM056TF
30	17	28.8	797	9	CNS02XJY	AL218455 Tetradon
31	17	28.8	800	5	AG215582	AG215582 Drosophila
32	17	28.8	930	5	BU185878	BU185878 AGENCOURT
33	17	28.8	950	6	CC129912	CC129912 NDJL.49015
34	17	28.8	966	2	BE314981	BE314981 601140460
35	17	28.8	983	9	CL111728	CL111728 ISB1-5512
36	17	28.8	1088	5	BO917972	BO917972 AGENCOURT
37	17	28.8	1106	9	CG410407	CG410407 AGTCS566X
38	17	28.8	1111	8	CC277811	CC277811 CH261-30K
39	17	28.8	1121	8	CC273227	CC273227 CH261-131
40	17	28.8	1152	5	BU230193	BU230193 603800301
41	16	27.1	163	1	AA745734	AA745734 NY93D01.8
42	16	27.1	183	6	CA365559	CA365559 642158 NC
43	16	27.1	205	2	BF882752	BF882752 QV3-ET017
44	16	27.1	231	9	AG251477	AG251477 Lotub cor
45	16	27.1	232	6	CD157665	CD157665 MLI-0053T

ALIGNMENTS

RESULT 1
LOCUS CL361142
DEFINITION RPTC14_344U4.f RPTC14_44 Sus scrofa genomic clone RPTC14_344U4,
genomic survey sequence.
ACCESSION CL361142
VERSION CL361142.1 GI:51413112
KEYWORDS GSS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

REFERENCE

1 (bases 1 to 445)
Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marrou,B.M.,
Mammalia; Butneria; Cetartiodactyla; Suidae; Sus.

AUTHORS

Beever,J.E. and Schook,L.B.

Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
Through Comparative Genomics

JOURNAL

Unpublished (2004)
Other GSSs: RPTC14_344U4.r

Contact: Lawrence B. Schook
Department of Animal Sciences

University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA

Tel: 217 265 5926
Fax: 217 244 5617

Email: schook@uiuc.edu

Clones are derived from the porcine BAC library RPTC14-44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@chori.org).

Clones may be purchased from BACPAC Resources
(http://BACPACresources.chori.org). This work was undertaken as part

of the International Swine Genome Sequencing Consortium by
University of Illinois at Urbana Champaign, USA with funds provided

by grant No. AG2002-34480-11828 from USDA-CSREES and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing

Initiative)

Plate: 344 row: J column: 4

Seq primer: 17

Class: BAC ends.

Location/Qualifiers

1..445

/organism="Sus scrofa"

/mol_type="Genomic DNA"

/strain="four pigs (bred: 37.5% Yorks Landrace and 25%
Meishan)"

/db_xref="taxon:9823"

/clone="RPTC14_344U4"

FEATURES

source

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-22

Perfect score: 60

Sequence: 1 CATCTACAGTTATTCCTTG.....TATAATACATGATCAT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2359870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseq1980s:*\n2: Geneseq1990s:*\n3: Geneseq2000s:*\n4: Geneseq2001s:*\n5: Geneseq2001bs:*\n6: Geneseq2002as:*\n7: Geneseq2002bs:*\n8: Geneseq2003as:*\n9: Geneseq2003bs:*\n10: Geneseq2003cs:*\n11: Geneseq2003ds:*\n12: Geneseq2004as:*\n13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23194
2	20	33.3	59	2	AAV23306
3	19	31.7	58	2	AAV23186
4	17	28.3	58	6	ABK98625
5	17	28.3	58	9	ACD13876
6	17	28.3	59	2	AAV23193
7	17	28.3	60	6	ABK98622
8	17	28.3	60	9	ACD13873
9	17	28.3	64	6	ABK98604
10	17	28.3	64	6	ABK98605
11	17	28.3	65	6	ABK98605
12	17	28.3	65	9	ACD13856
13	17	28.3	93	6	ABK98615
14	17	28.3	93	6	ABK98615
15	17	28.3	95	6	ABK98612
16	17	28.3	95	9	ACD13863
17	17	28.3	115	6	ABK98626
18	17	28.3	115	6	ABK98626
19	17	28.3	150	6	ABK98616
20	17	28.3	150	9	ACD13867

C	21	17	28.3	476	4	AAK81851
C	22	17	28.3	2300	4	AAK123400
23	17	28.3	9417	6	ABK98632	
24	17	28.3	9417	9	ACD13883	
25	17	28.3	9437	6	ABK98601	
26	17	28.3	9437	9	ACD13852	
27	17	28.3	9484	6	ABK98633	
28	17	28.3	9484	9	ACD13884	
29	17	28.3	10929	6	ABK98591	
30	17	28.3	10929	9	ACD13842	
31	17	28.3	12739	6	ABK98592	
32	17	28.3	12739	9	ACD13843	
33	17	28.3	110000	6	ABA90521_00	
34	16	26.7	59	2	AAV23210	
35	16	26.7	60	2	AAV23188	
36	16	26.7	60	2	AAV23200	
37	16	26.7	60	2	AAV23212	
38	16	26.7	60	2	AAV23199	
C	39	16	26.7	531	10	ADD17215
C	40	16	26.7	531	10	ADD17215
41	16	26.7	755	6	ABN99221	
42	16	26.7	894	10	ADR62321	
43	16	26.7	988	10	ABE56168	
44	16	26.7	1062	6	ABZ13652	
45	16	26.7	1062	6	ADG87654	

ALIGNMENT

ALIGNMENTS

RESULT 1

ID AAV23194 standard; DNA; 60 BP.

AC AAV23194;

DT 28-JUL-1998 (first entry)

XX Lactococcus lactis constitutional promoter Cp26.

XX Lactococcus lactis; constitutional promoter; optimise; spacer;

KW artificial promoter library; gene expression; de.

XX Synthetic.

OS Lactococcus lactis.

FT Key

FT promoter

PN WO9807846-A1.

PD 26-FEB-1998.

PF 25-AUG-1997; 97WO-DK00342.

PR 23-AUG-1996; 96DK-00000886.

XX (JENS/) JENSEN P R.

XX Jensen PR, Hammer K;

XX WPI; 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter

XX PT sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.

XX Claim 28; Page 51; 89pp; English.

CC This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds

5668.525 Million cell updates/sec

Title: US-09-242-657B-22

Sequence: 1 CATTCTACAGTTTATTCCTTG.....TATAATACTATACATGCAT 60

Scoring table: OLIGO_NUC

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Post-processing: Listing first 45 summaries

Database : GenEmbl : *

```

1: GenEmbl: *
2:   gb_ba: *
3:   gb_htg: *
4:   gb_in: *
5:   gb_om: *
6:   gb_ov: *
7:   gb_pat: *
8:   gb_ph: *
9:   gb_pl: *
10:  gb_pr: *
11:  gb_ro: *
12:  gb_sts: *
13:  gb_sy: *
14:  gb_un: *
   gb_vi: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	6	BD006953	BD06553 Artificialia
2	22	36.7	188406	10	AL713914	AL713914 Mouse DNA
3	21	35.0	130530	9	AF131217	AF131217 Homo sapi
4	21	35.0	134614	9	BS000225	BS000225 Pan trogl
5	21	35.0	142215	9	BS000165	BS000165 Pan trogl
6	21	35.3	340000	9	HS21C047	AL163347 Homo sapi
7	20	33.3	59	6	BD006965	BD006965 Artificialia
8	19	31.7	58	6	BD006945	BD006945 Artificialia
9	19	31.7	137230	9	AC027973	AC027973 Homo sapi
10	19	31.7	161598	2	AC127513	AC127513 Homo sapi
11	19	31.7	174447	2	AC142552	AC142552 Danto rer
12	19	31.7	182160	2	AC142550	AC142550 Danto rer
13	19	31.7	195778	5	BX088566	AX088566 Zebrafish
14	19	31.7	199182	9	AC005562	AC005562 Homo sapi
15	19	31.7	200704	9	AC099781	AC099781 Homo sapi
16	19	31.7	212934	9	AC138207	AC138207 Homo sapi
17	19	31.7	225287	2	AC127530	AC127530 Homo sapi
18	19	31.7	245174	2	AC125806	AC125806 Rattus no
19	19	31.7	253236	2	BX950188	BD50188 Danto rer

C	20	18	30.0	4.2163	3	CEP49C5	28154	Caenorhabdi
C	21	18	30.0	100845	5	EX470240	EX470240	Zebrafish
C	22	18	30.0	147187	10	ALB07393	ALB07393	Mouse DNA
C	23	18	30.0	148396	9	AL157400	AL157400	Human DNA
C	24	18	30.0	158758	9	AC093591	AC093591	Homo sapi
C	25	18	30.0	160654	2	CR788241	CR788241	Danio rer
C	26	18	30.0	163933	5	AL2929559	AL2929559	Zebrafish
C	27	18	30.0	177219	2	AC021286	AC021286	Homo sapi
C	28	18	30.0	188357	9	AL159966	AL159966	Homo sapi
C	29	18	30.0	190837	9	AF129075	AF129075	Human DNA
C	30	18	30.0	192836	9	BS000157	BS000157	Homo sapi
C	31	18	30.0	195748	10	ALB71869	ALB71869	pan trogl
C	32	18	30.0	199432	2	AC107233	AC107233	Mouse DNA
C	33	18	30.0	204257	2	AC104293	AC104293	Mus muscu
C	34	18	30.0	211935	2	AC135941	AC135941	Rattus no
C	35	18	30.0	212042	2	AC130593	AC130593	Rattus no
C	36	18	30.0	218680	9	BS000171	BS000171	Rattus no
C	37	18	30.0	219157	2	AC128726	AC128726	pan trogl
C	38	18	30.0	228633	5	EX649600	EX649600	Rattus no
C	39	18	30.0	232964	2	AC015501	AC015501	Zebrafish
C	40	18	30.0	233713	2	AC098341	AC098341	Homo sapi
C	41	18	30.0	242452	2	AC093960	AC093960	Rattus no
C	42	18	30.0	285659	2	AC128227	AC128227	Rattus no
C	43	18	30.0	340000	9	HS21C049	HS21C049	Homo sapi
C	44	17	28.3	59	6	BD006952	BD006952	Artificia
C	45	17	28.3	67	6	AR055664	AR055664	Sequence

ALIGNMENTS

LOCUS	DEFINITION
BD006953	60 bp DNA linear PAT 31-JAN-2002
BD006953	Artificial promoter libraries for selected organisms and promoters derived from such libraries.

ACCESSION	BD006953
VERSION	BD006953.1
KEYWORDS	GI:18635324
	JP 2001503249-A/22.

ORGANISM

00000000

AUTHORS

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843

PETER RUDAR JANSSEN

FEATURE

OPTGTR

Query

Best
Match

22

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-22

Perfect score: 60
Sequence: 1 CATCTACTAGTTTATCTTG.....TATAATACTATACATCAT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST:*

1: gb_est1:*\n2: gb_est2:*\n3: gb_est3:*\n4: gb_est4:*\n5: gb_est5:*\n6: gb_est6:*\n7: gb_est7:*\n8: gb_est8:*\n9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	437	7	NS8914
2	19	31.7	551	8	AQ693640
3	19	31.7	729	6	CD110702
4	19	31.7	961	4	BG698770
5	19	31.7	1111	8	CC277811
6	18	30.0	469	1	AJ637534
7	18	30.0	557	1	AJ636028
8	18	30.0	656	9	EX127652
9	18	30.0	728	6	CA847206
10	18	30.0	746	9	AZ976321
11	18	30.0	892	4	BR977586
12	18	30.0	1065	8	CC298357
13	18	30.0	1254	9	BK660524
14	17	28.3	307	9	U27999
15	17	28.3	310	1	AJ796482
16	17	28.3	330	9	CR307025
17	17	28.3	365	8	AQ284199
18	17	28.3	369	8	AZ709398
19	17	28.3	390	5	BQ161611
20	17	28.3	396	5	AQ103157
21	17	28.3	397	5	BQ635080
22	17	28.3	397	5	BQ655889
23	17	28.3	397	5	BQ655889
24	17	28.3	397	5	BQ655889

C 25	17	28.3	399	1	AI872271
C 26	17	28.3	407	2	AM015492
C 27	17	28.3	450	8	AQ884083
C 28	17	28.3	458	1	AI920068
C 29	17	28.3	479	1	AJ685409
C 30	17	28.3	526	4	BM427762
C 31	17	28.3	564	9	CC772164
C 32	17	28.3	596	9	EX215823
C 33	17	28.3	600	7	CO349954
C 34	17	28.3	609	4	BM135412
C 35	17	28.3	627	8	AO233914
C 36	17	28.3	629	8	AZ741223
C 37	17	28.3	635	8	BH003371
C 38	17	28.3	654	4	BG888975
C 39	17	28.3	668	8	AZ624988
C 40	17	28.3	670	9	AG108977
C 41	17	28.3	671	8	BZ001668
C 42	17	28.3	676	9	CE237580
C 43	17	28.3	681	9	CC527143
C 44	17	28.3	683	7	CN625364
C 45	17	28.3	712	8	AZ738674

ALIGNMENTS

RESULT 1
LOCUS NS8914
DEFINITION Yy61e12.81 Soares multiple sclerosis_2bhmSP Homo sapiens CDNA
clone IMAGE:278062 3', mRNA sequence.
NS8914
NS8914.1 GI:1202804
EST.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 437)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rickin, L., Rohlfing, T., Soares, M., Tan, P.,
Trevaskis, E., Waterston, R., Williamson, A., Woldmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1900
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mod.RBGA+RT
High quality sequence stop: 1.
Location/Qualifiers
1. 437
/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="GDB:3896438"
/db_xref="taxon:9606"
/clone="IMAGE:278062"
/sex="male"
/tissue_type="multiple sclerosis lesions"
/dev_stage="Age 46"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares multiple sclerosis_2bhmSP"
/note="Vector: pTR3D (Pharmacia) with a modified
polylinker V type: phagemid; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5',
TTTACCATCTGAGTGGAGCGCCGCGCATTTTCTTTTCTTTT 3'],

TITLE
JOURNAL
COMMENT

FEATURES

source

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657b-23

Perfect score: 60

Sequence: 1 CATTGGGCGCTTTTCTCTG.....TATTAATATAGTACTGTT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2 AAV23195	Aav23195 Lactococc
2	18	30.0	477	6 ABN60951	Abn60951 Human can
3	17	28.3	350	3 AAC30261	Aac30261 Human sec
4	17	28.3	2918	11 ADM74716	Adm74716 B. Churth
5	17	28.3	6528	3 AAZ49723	Aaz49723 Staphyloc
6	17	28.3	6528	10 ACF36709	Acf36709 S. stimula
7	17	28.3	110000	8 ABO83210	Abg83210 Human tra
8	17	28.3	192427	10 ADL13825	Adl13825 Osteoarth
9	16	26.7	265	2 AAQ23880	Aaq23880 PRT1820
10	16	26.7	265	2 AAQ46205	Aaq46205 Promoter
11	16	26.7	265	2 AAT31875	Aat31875 Promoter
12	16	26.7	399	8 ABX54585	Abx54585 Bovine ES
13	16	26.7	583	5 ABV53893	Abv53893 Human pro
14	16	26.7	814	6 ABK95277	Abk95277 Human pro
15	16	26.7	1203	6 AAD47003	Aad47003 Plasmodiu
16	16	26.7	1205	6 ABK95278	Abk95278 Human pro
17	16	26.7	1494	8 ACA35365	Aca35365 Prokaryot
18	16	26.7	1596	11 ACH96815	Ach96815 Klebsiell
19	16	26.7	2382	4 ABL29743	Ab129743 Drosophil
20	16	26.7	5274	4 ABL15583	Ab115583 Drosophil

C	21	16	26.7	8785	4 ABL15582	Ab115582 Drosophil
C	22	16	26.7	9455	4 ABL10283	Ab110283 Drosophil
C	23	16	26.7	12045	4 AAL07066	Aal07066 Human rep
C	24	16	26.7	13036	4 ABL10282	Ab110282 Drosophil
C	25	16	26.7	14428	4 ABL29742	Ab129742 Drosophil
C	26	16	26.7	71138	12 AD013830	Ad013830 dbv gene
C	27	16	26.7	110000	6 ABA90521_02	Continuation (3 of
C	28	15	25.0	25	9 ACT20369	Act20369 Human mtc
C	29	15	25.0	59	2 AAV23209	Aav23209 Lactococc
C	30	15	25.0	60	2 AAV23213	Aav23213 Lactococc
C	31	15	25.0	115	6 ABK98626	Abk98626 L. lactis
C	32	15	25.0	115	9 ACD13877	Act13877 L. lactis
C	33	15	25.0	150	6 ABK98616	Abk98616 L. lactis
C	34	15	25.0	150	9 ACD13867	Act13867 L. lactis
C	35	15	25.0	201	13 AD042111	Ad042111 Myocardia
C	36	15	25.0	207	6 ABK98627	Abk98627 L. lactis
C	37	15	25.0	207	9 ACD13878	Act13878 L. lactis
C	38	15	25.0	242	6 ABK98617	Abk98617 L. lactis
C	39	15	25.0	242	9 ACD13868	Act13868 L. lactis
C	40	15	25.0	272	3 AAA42899	Aaa42899 Human sec
C	41	15	25.0	327	11 ABD06404	Abd06404 Pseudomon
C	42	15	25.0	363	6 ABK76445	Abk76445 Bacillus
C	43	15	25.0	395	10 ADB51088	Adb51088 Primary r
C	44	15	25.0	447	4 AAK59279	Aak59279 Human ltm
C	45	15	25.0	500	3 AAC95157	Aac95157 Cat flea

ALIGNMENTS

RESULT 1
AAV23195
ID AAV23195 standard; DNA; 60 BP.

XX AAV23195;

DT 28-JUL-1998 (first entry)

XX Lactococcus lactis constitutional promoter Cp28.

XX Lactococcus lactis; constitutional promoter; optimise; spacer;

XX Lactococcus lactis.

OS Lactococcus lactis.

XX Key Location/Qualifiers

FT promoter 4.60 /tag= a

FT /standard_name= "Constitutional promoter"

PN WO9807846-A1.

XX 26-FEB-1998.

PF 25-AUG-1997; 97WO-DK000342.

PR 23-AUG-1996; 96DK-00000886.

PA (JENSEN) JENSEN P. R.

XX Jensen PR, Hammer K;

PI WPI; 1998-179062/16.

DR New artificial promoter libraries - containing consensus promoter

PS sequences and variable spacers, used to generate promoters for optimising

CC expression of genes.

This is a Lactococcus lactis constitutional promoter sequence used in the construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds

(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-23

Perfect score: 60
Sequence: 1 CATGGGGCGCTTATCTCTG.....TATATATATATAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4708233 segs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Geneml:*

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_rts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006954	BD006954 Artificial
2	19	31.7	143168	AP003765	AP003765 Oryza sat
3	19	31.7	146680	AP004671	AP004671 Oryza sat
4	19	31.7	184598	AL139406	AL139406 Human DNA
5	19	31.7	190357	AC012259	AC012259 Homo sapi
6	18	30.0	477	AX385990	AX385990 Sequence
7	18	30.0	97751	AC084257	AC084257 Homo sapi
8	18	30.0	123883	AL845278	AL845278 Mouse DNA
9	18	30.0	143819	AC137765	AC137765 Mus muscu
10	18	30.0	155821	AC027043	AC027043 Homo sapi
11	18	30.0	160673	CNS080CDR	CNS080CDR Oryza sat
12	18	30.0	169262	AC090569	AC090569 Homo sapi
13	18	30.0	174751	AC009137	AC009137 Homo sapi
14	18	30.0	175038	AC150652	AC150652 Bos tauru
15	18	30.0	177584	AC023116	AC023116 Homo sapi
16	18	30.0	193761	AC022273	AC022273 Homo sapi
17	18	30.0	200022	AC137654	AC137654 Bos tauru
18	18	30.0	241430	AC115560	AC115560 Rattus no
19	18	30.0	243483	AC116256	AC116256 Rattus no

20	18	30.0	246731	2	AC097311	AC097311 Rattus no
21	18	30.0	250050	1	AL627273	AL627273 Salmonella
22	18	30.0	250828	2	AC097070	AC097070 Rattus no
23	18	30.0	300928	1	AE016836	AE016836 Salmonella
24	17	28.3	350	6	AX918473	AX918473 Sequence
25	17	28.3	350	6	BD054006	BD054006 Sequence
26	17	28.3	1250	14	AF389459	AF389459 Lymantria
27	17	28.3	2920	1	AF211190	AF211190 Bacillus
28	17	28.3	3214	8	AY177888	AY177888 Sorghum b
29	17	28.3	6451	6	AX828352	AX828352 Sequence
30	17	28.3	6528	1	SSU66883	SSU66883 Stephylococ
31	17	28.3	49940	5	BX649373	BX649373 Zebrafish
32	17	28.3	56345	10	AL672085	AL672085 Mouse DNA
33	17	28.3	76568	3	MREV	AF38053 Monosiga
34	17	28.3	91573	2	AC012372	AC012372 Homo sapi
35	17	28.3	96925	2	AC151866	AC151866 Ornithoch
36	17	28.3	99268	9	AL589948	AL589948 Human DNA
37	17	28.3	109306	9	AC093529	AC093529 Homo sapi
38	17	28.3	110000	1	BX950851	Continuation (6 of
39	17	28.3	139589	9	AC010470	AC010470 Homo sapi
40	17	28.3	143035	9	AC010082	AC010082 Homo sapi
41	17	28.3	152341	9	AL137848	AL137848 Human DNA
42	17	28.3	153089	9	AC146510	AC146510 Pan trogl
43	17	28.3	160039	9	AC094109	AC094109 Homo sapi
44	17	28.3	160039	9	AC068323	AC068323 Homo sapi
45	17	28.3	165967	9	AC010349	AC010349 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS BD006954
DEFINITION Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION BD006954.1 GI:18635325
VERSION JP 2001503249-A/23.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE 1 (bases 1 to 60)
AUTHORS Hammer, K. and Janssen, P.R.
TITLE Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL Patent: JP 2001503249-A 23 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT OS Lactococcus lactis
PN JP 2001503249-A/23
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter Location/Qualifiers

FEATURES
SOURCE Location/Qualifiers
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.3e-24;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGGGGCGCTTATCTGACGACGACGACGCTGTATATATATATAGTACTGTT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-23

Perfect score: 60

Sequence: 1 CATGCGGCGCTTATCTCTG.....TATATATATATAGACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	19	31.7	484	8	AZ198221 SP 1037 A
2	19	31.7	580	2	BE713377 MRI-HR071
3	19	31.7	594	8	AQ475793 CITBI-E1-
4	18	30.0	122	8	CC096838 CSU-K34.1
5	18	30.0	536	9	AG252334 locus cor
6	18	30.0	667	9	CL200124 ZMMB022
7	18	30.0	760	6	CD376536 PTM01704
8	18	30.0	803	4	CL638498 CH243-2D2
9	18	30.0	1028	4	BE844674 1024007D0
10	17	28.3	123	6	CB925588 ABA1_22_F
11	17	28.3	298	8	AQ021759 CIT-HSP-2
12	17	28.3	407	9	BX123136 Dario rer
13	17	28.3	452	9	CE484970 tigr-gss-
14	17	28.3	483	9	CR396152
15	17	28.3	528	6	CR3685
16	17	28.3	534	6	CB927620 ABA1_27_D
17	17	28.3	548	6	BQ313337 RCO-BN032
18	17	28.3	557	6	BW200077 BM200077
19	17	28.3	574	6	CB926880 ABA1_11_A
20	17	28.3	588	7	CK509930 rswdd_00
21	17	28.3	589	6	CB926018 ABA1_31_D
22	17	28.3	595	6	BW201140 BM201140
23	17	28.3	598	6	CB928453 ABA1_16_C
24	17	28.3	606	6	CB617307 3529_1_72

C 25	17	28.3	617	6	CB924989 ABA1_29_D
C 26	17	28.3	649	6	CD225439 CCCT1_3_D0
C 27	17	28.3	650	6	CB437031 684465_MA
C 28	17	28.3	651	6	CD219529 CCCT1_57_A
C 29	17	28.3	660	8	AQ714761 HS_5423_B
C 30	17	28.3	667	8	BH201111 Sml-55KI.
C 31	17	28.3	679	7	CF765710 CR5001452
C 32	17	28.3	690	5	BW448653 BM448653
C 33	17	28.3	694	8	BH983021 Odf77E01.
C 34	17	28.3	737	6	CD211308 HSI_59_F1
C 35	17	28.3	752	1	AJ817583 AJ817583
C 36	17	28.3	764	9	BX220289 Dario rer
C 37	17	28.3	774	9	CG212895 CGVD231TV
C 38	17	28.3	776	3	CR709541 Tetradon
C 39	17	28.3	783	5	BU451678 BX209903
C 40	17	28.3	784	9	BX209903 Dario rer
C 41	17	28.3	800	7	CN022324 AGENCOURT
C 42	17	28.3	802	5	BU401456 603485569
C 43	17	28.3	806	8	BZ679699 PUBHRT1TD
C 44	17	28.3	806	9	AG522944 MNB muscu
C 45	17	28.3	890	9	CL902521 CSNC1896

ALIGNMENTS

RESULT 1
AZ198221
LOCUS
DEFINITION
SP 1037 A1 F07 SPEE Strongylocentrotus purpuratus, purple sea urchin sperm genomic BAC library Strongylocentrotus purpuratus genomic clone Plate=1037 Col=13 Row=K, genomic survey sequence.
AZ198221 GI:8392046

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
Strongylocentrotus purpuratus
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidae; Euechinoidae; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 484)

REFERENCE
AUTHORS
Cameron, R.A., Mahiras, G., Rast, J.P., Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Ponsky, A.J., Livingston, B.T., Wray, G.A., Btenschn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H. and Hood, L.

TITLE
A sea urchin genome project: Sequence scan, virtual map, and additional resources
JOURNAL
PROC. NATL. ACAD. SCI. U.S.A. 97 (17), 9514-9518 (2000)
MEDLINE
PUBMED
10920195

COMMENT
Contact: Cameron, R.A., Davidson, E.H., Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 1037 row: K column: 13
Seq primer: SP6
Clase: BAC ends
High quality sequence stop: 484.

FEATURES
Location/Qualifiers
1..484

source
/organism="Strongylocentrotus purpuratus"
/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone_plate="1037 Col=13 Row=K"
/clone_lib="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACE3.6; BAC Clones in E-Coli DH10B"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 195.613 Seconds
(without alignments)
1785,485 Million cell updates/sec

Title: US-09-242-657B-24

Perfect score: 59
Sequence: 1 CATCGGTAGTATTCTTGA.....TATTAATAGTACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseq19808:*
2: Geneseq19908:*
3: Geneseq20008:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	2	AAV23196 Lactococc
2	20	33.9	60	2	AAV23211 Lactococc
3	17	28.8	844	4	AAH08445 Human CDN
4	17	28.8	12194	4	AA828921 Human imm
5	17	28.8	12194	10	ADB31762 Human nov
6	16	27.1	210	6	ABN93179 Streptyl
7	16	27.1	210	13	ADSO4092 Streptyl
8	16	27.1	253	6	ABN21544 Human ORF
9	16	27.1	447	4	ABA08639 Human sec
10	16	27.1	447	9	ACH25465 Human adu
11	16	27.1	840	5	AA891304 DNA encod
12	16	27.1	1481	6	AAAD4639 Human amy
13	16	27.1	1481	6	AAAD4686 Human amy
14	16	27.1	2989	4	ABL07272 Drosophi
15	16	27.1	10772	4	AAK6425 Human imm
16	16	27.1	93329	13	ABD33597 Murine ca
17	16	27.1	110000	13	ABD32804-2 Continuation (3 of
18	16	27.1	134442	13	ABD32824 Mouse can
19	15	25.4	201	13	ADQ42111 Myocardia
20	15	25.4	388	9	ACH21096 Human adu

C 21	15	25.4	432	8	ABX49842	Abx49842 Bovine BS
C 22	15	25.4	439	4	AAH12334	Aah12334 Human CDN
C 23	15	25.4	443	5	AA875470	Aa875470 DNA encod
C 24	15	25.4	491	13	ACN56756	Acn56756 Cotton gy
C 25	15	25.4	506	6	ABN96059	Abn96059 Gene #256
C 26	15	25.4	527	13	ACN47081	Acn47081 Cotton pr
C 27	15	25.4	547	13	ACN48829	Acn48829 Cotton pr
C 28	15	25.4	549	6	ABK75169	Abk75169 Bacillus
C 29	15	25.4	565	13	ACN48754	Acn48754 Cotton pr
C 30	15	25.4	570	13	ACN47008	Acn47008 Cotton pr
C 31	15	25.4	590	4	AA860518	Aa860518 Human can
C 32	15	25.4	606	10	ADG31898	Adg31898 Chicken m
C 33	15	25.4	610	13	ACN56668	Acn56668 Cotton gy
C 34	15	25.4	642	5	AAH52231	Aah52231 Human AFP
C 35	15	25.4	688	11	ADT95691	Adt95691 Colon can
C 36	15	25.4	738	15	AAI96607	Aai96607 Human neu
C 37	15	25.4	759	6	ABK53355	Abk53355 Human col
C 38	15	25.4	774	13	AD849475	Ad849475 Bacterial
C 39	15	25.4	775	4	AA860748	Aa860748 Human can
C 40	15	25.4	804	2	AAQ76186	Aaq76186 Calcium s
C 41	15	25.4	804	2	AAT28771	Aat28771 Human cal
C 42	15	25.4	804	2	AAV05992	Aav05992 Human cal
C 43	15	25.4	813	6	ABQ93347	Abq93347 Human cod
C 44	15	25.4	844	12	ADP13390	Adp13390 Renal cel
C 45	15	25.4	862	4	AA860745	Aa860745 Human can

ALIGNMENTS

RESULT 1
AAV23196 standard; DNA; 59 BP.

AC AAV23196;

DT 28-JUL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cp29.

KW Lactococcus lactis; constitutional promoter; optimise; spacer;

KM artificial promoter library; gene expression; ds.

OS Synthetic.

OS Lactococcus lactis.

FT Key Location/Qualifiers

FT promoter 4..59

FT FT /tag= a

PN MO9807846-A1. /standard_name= "Constitutional promoter"

PD 26-FEB-1998.

XX 25-AUG-1997; 97WO-DK000342.

XX 23-AUG-1996; 96DK-00000886.

XX (JENS/) JENSEN P R.

XX Jensen PR, Hammer K;

XX WPI; 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter

XX sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.
Claim 28; Page 52; 89pp; English.
This is a Lactococcus lactis constitutional promoter sequence used in the construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 504.339 Seconds

(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657b-24

Perfect score: 59

Sequence: 1 CATCGTAAGTATTTCTTGA.....TATTAATAGTACTGTT 59

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenBank: *
1: gb_ba: *
2: gb_ptg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_dl: *
9: gb_dr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	59	100.0	59	BD006955
2	59	33.9	60	BD006970
3	32.2	84553	4	AC123538
4	32.2	112662	2	AC104453
5	32.2	252953	2	AC133431
6	32.2	281573	2	AC125743
7	30.5	1328	8	AF377771
8	30.5	1384	8	AF044292
9	30.5	1387	8	AF328413
10	30.5	1387	8	AF328417
11	30.5	1387	8	AF328417
12	30.5	1387	8	AF328422
13	30.5	1389	8	AF044289
14	30.5	2082	8	MZEMAYADH
15	30.5	47167	2	AC048373
16	30.5	152631	2	AC048373
17	30.5	177490	9	AP000760
18	30.5	195409	9	AP004609
19	30.5	232385	2	AC133770

20	18	30.5	241573	2	AC111817	AC111817 Rattus no
21	18	30.5	242109	2	AC128430	AC128430 Rattus no
22	18	30.5	244847	2	AP002357	AP002357 Homo sapi
23	17	28.8	187	11	G53465	G53465 SHGC-85254
24	17	28.8	844	6	BD150437	BD150437 Primer fo
25	17	28.8	844	6	AX870375	AX870375 Sequence
26	17	28.8	32715	3	CHRS4P12	Z81548 Caenorhabdi
27	17	28.8	71641	9	AC097380	AC097380 Homo sapi
28	17	28.8	72143	9	AC131094	AC131094 Homo sapi
29	17	28.8	97736	9	AC093845	AC093845 Homo sapi
30	17	28.8	120873	9	AL445986	AL445986 Human DNA
31	17	28.8	140415	5	CR392037	CR392037 Zebrafish
32	17	28.8	140780	8	OSJN00041	AL666630 Oryza sat
33	17	28.8	145715	2	AC117018	AC117018 Rattus no
34	17	28.8	151959	9	AP002089	AP002089 Homo sapi
35	17	28.8	155544	8	AP005799	AP005799 Oryza sat
36	17	28.8	156054	8	AB023482	AB023482 Oryza sat
37	17	28.8	157905	8	AP005495	AP005495 Oryza sat
38	17	28.8	161582	2	AC020628	AC020628 Homo sapi
39	17	28.8	163201	9	AC093814	AC093814 Homo sapi
40	17	28.8	165549	9	AL355498	AL355498 Human DNA
41	17	28.8	165570	2	AC015765	AC015765 Homo sapi
42	17	28.8	168231	2	AC119900	AC119900 Mus muscu
43	17	28.8	168813	9	AC013717	AC013717 Homo sapi
44	17	28.8	170695	2	AC068545	AC068545 Homo sapi
45	17	28.8	171703	2	AC023504	AC023504 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS BD006955
DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION BD006955.1 GI:18635326
VERSION JP 2001503249-A/24.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE 1 (bases 1 to 59)
AUTHORS Hammer,K. and Janssen,P.R.
TITLE Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL Patent: JP 2001503249-A 24 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Lactococcus lactis
PN JP 2001503249-A/24
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
FH Key
FT promoter (4)..(59).
Location/Qualifiers

FEATURES

1..59
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 59; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATCGTAAGTATTTCTTGAATCTGAGGGGAGCGTGTATTAATAGTACTGTT 59

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 / Search time 1430.81 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-24

Perfect score: 59
Sequence: 1 CATCGGTAGTATCTTGA.....TATATACGAGTACTT 59

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST:

1: gb_est1:*
2: gb_est2:*
3: gb_hrc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18	30.5	836 8 BH202020	Sml-55E8.
2	18	30.5	1066 8 BM909534	BM909534 AGNCORT
3	17	28.8	187 8 AQ082883	AQ082883 RPT11-55
4	17	28.8	454 8 AQ061689	AQ061689 HS_5208_A
5	17	28.8	510 8 AQ057441	AQ057441 HS_2081_A
6	17	28.8	514 9 CE091165	CE091165 CIGR-g88
7	17	28.8	523 8 AQ058173	AQ058173 RPT1-11-4
8	17	28.8	540 4 B2699937	B2699937 B689937
9	17	28.8	603 8 A2228847	A2228847 RPT1-23-9
10	17	28.8	630 8 B2495477	B2495477 BONNT17TF
11	17	28.8	637 9 AG102502	AG102502 Pan trogl
12	17	28.8	652 6 CB502131	CB502131 ssalmg950
13	17	28.8	680 8 AQ056735	AQ056735 HS_2109_A
14	17	28.8	682 8 B2881889	B2881889 CH240_253
15	17	28.8	730 9 CE734287	CE734287 CIGR-g88
16	17	28.8	769 8 BH555820	BH555820 B0HNY57TF
17	17	28.8	779 8 AG025633	AG025633 Homo sapi
18	17	28.8	838 8 BH461589	BH461589 B0GBS20TF
19	17	28.8	848 8 BZ264093	BZ264093 CH230-521
20	17	28.8	909 7 CC379202	CC379202 PUTHH21TD
21	17	28.8	919 7 CV289721	CV289721 acf01-6ms
22	17	28.8	1059 9 CL985629	CL985629 ZMMBH000
23	17	28.8	1094 8 CC209486	CC209486 CH261-72P
24	16	27.1	113 8 AQ413156	AQ413156 RPT1-11-1

25	16	27.1	209 2 AW917904	AW917904 EST349208
26	16	27.1	283 9 CC770461	CC770461 CH240_5K1
27	16	27.1	293 9 CR270150	CR270150 RPTW4_8
28	16	27.1	304 7 CO345256	CO345256 EP21710.3
29	16	27.1	332 2 AM291622	AM291622 UI-H-B12
30	16	27.1	335 7 D64270	D64270 CELK048B3R
31	16	27.1	342 6 CB177247	CB177247 BPT0302.Y
32	16	27.1	345 8 BH030309	BH030309 RPT1-24-2
33	16	27.1	349 8 BZ845899	BZ845899 CH240_216
34	16	27.1	353 9 CC773736	CC773736 CH240_33A
35	16	27.1	358 9 CG985885	CG985885 CH240_156
36	16	27.1	365 7 T01384	T01384 WEST02105_E
37	16	27.1	367 5 BP703811	BP703811 BPT03811
38	16	27.1	368 9 CC576134	CC576134 CH240_454
39	16	27.1	372 7 D33019	D33019 CELK025A7R
40	16	27.1	373 9 CC768851	CC768851 CH240_68B
41	16	27.1	376 6 CB069987	CB069987 1832a08.Y
42	16	27.1	376 9 CC474817	CC474817 CH240_300
43	16	27.1	383 5 BM956464	BM956464 C6aC43551
44	16	27.1	385 6 CB066870	CB066870 iQ32e08.x
45	16	27.1	388 8 AQ192402	AQ192402 HS_2251_A

ALIGNMENTS

RESULT 1
BH202020
LOCUS
DEFINITION
Sml-55E8, TR Sml Schistosoma mansoni genomic clone Sml-55E8, genomic survey sequence.
ACCESSION
BH202020
VERSION
BH202020.1
KEYWORDS
GSS.
SOURCE
Schistosoma mansoni
ORGANISM
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 836)
Shetty, J., Simpson, A., Malek, J., Koo, H., Loverde, P.T. and El-Sayed, N.M.
Use of end sequences from Schistosoma mansoni (Puerto Rico strain) Sml BAC library for gene discovery and map construction
Unpublished (2001)
Other_GSSes: Sml-55E8.TF
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
Clones are derived from the Schistosoma mansoni (Puerto Rico) Sml BAC library. For clone availability, please contact Dr. Najib El-Sayed at tigr (nelsayed@tigr.org) or Dr. Phillip Loverde at (loverde@buffalo.edu)
Seq primer: M13 Rev
Class: BAC ends.

FEATURES

source
1. 836
/organism="Schistosoma mansoni"
/mol_type="genomic DNA"
/strain="Puerto Rico"
/db_xref="taxon:6183"
/clone="Sml-55E8"
/clone_lib="Sml"
/note="Vector: pBelBAC11, site 1: Hin dIII; Constructed in the laboratory of Dr. Denis Le Paslier at the Fondation Jean Dausset, CEPH, Paris, France. Briefly, Schistosoma mansoni agarose embedded DNA was partially digested with Hin dIII. High molecular weight fragments were ligated in pBelBAC11 digested with Hin dIII. The average insert size

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-25

Perfect score: 60
Sequence: 1 CATCTGTACTTTATTTCTTG.....TATTAATGAGAGTACTGTT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 segs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseq19808:*
2: geneseq19908:*
3: geneseq20008:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23197 Lactococc
2	18	30.0	58	2	AAV23184 Lactococc
3	18	30.0	110000	12	ADQ97331_4 Continuation (5 of
4	17	28.3	59	2	AAV23209 Lactococc
5	17	28.3	60	2	AAV23211 Lactococc
6	17	28.3	60	2	AAV23182 Lactococc
7	17	28.3	60	2	AAV23214 Lactococc
8	17	28.3	60	2	AAV23212 Lactococc
9	17	28.3	60	2	AAV23205 Lactococc
10	17	28.3	60	2	AAV23205 Lactococc
11	17	28.3	100660	13	ABD32721_3 Continuation (4 of
12	16	26.7	60	2	AAV23190 Lactococc
13	16	26.7	60	2	AAV23188 Lactococc
14	16	26.7	60	2	AAV23308 Lactococc
15	16	26.7	60	2	AAV23213 Lactococc
16	16	26.7	60	2	AAV23178 Lactococc
17	16	26.7	60	2	AAV23180 Lactococc
18	16	26.7	207	6	ABK98627 L. lactis
19	16	26.7	242	6	ABK98617 L. lactis
20	16	26.7	242	9	ACD13868 L. lactis

ALIGNMENTS

C	21	16	26.7	542	6	ABK88552	Abk88552 DNA encod
C	22	16	26.7	542	6	ABK88553	Abk88553 Dog prost
C	23	16	26.7	542	12	AD052586	Ad052586 Dog COX-2
C	24	16	26.7	542	12	AD052584	Ad052584 Dog COX-2
C	25	16	26.7	651	3	AAFI3601	Aafi3601 Aspergill
C	26	16	26.7	1724	13	AD551253	Ad551253 Bacterial
C	27	16	26.7	1812	6	ABK88557	Abk88557 Dog prost
C	28	16	26.7	1812	6	ABK88556	Abk88556 Dog prost
C	29	16	26.7	1812	12	AD052590	Ad052590 Dog COX-2
C	30	16	26.7	1812	12	AD052591	Ad052591 Dog COX-2
C	31	16	26.7	2864	12	ADP50636	Adp50636 Murine DN
C	32	16	26.7	2928	10	ADP52709	Adp52709 Primary r
C	33	16	26.7	2928	12	ADP27274	Adp27274 Renal tox
C	34	16	26.7	3509	6	ABK88554	Abk88554 DNA encod
C	35	16	26.7	3509	6	ABK88555	Abk88555 Dog prost
C	36	16	26.7	3509	12	AD052587	Ad052587 Dog COX-2
C	37	16	26.7	3509	12	AD052589	Ad052589 Dog COX-2
C	38	16	26.7	3632	6	ABA94342	Ab94342 Canine cy
C	39	16	26.7	6763	6	ABO67068	Ab067068 Human ang
C	40	16	26.7	9566	6	ABK98634	Abk98634 Vector pr
C	41	16	26.7	9566	9	ACD13885	Acid13885 L. lactis
C	42	16	26.7	17656	6	AB555502	Ab555502 Sweetgum
C	43	16	26.7	33053	6	ABQ67005	Abq67005 Human ang
C	44	16	26.7	68940	2	AA57351	Aa57351 Human chr
C	45	16	26.7	110000	6	ABA90521_00	Ab90521 Genomic s

RESULT 1	AAV23197	standard; DNA; 60 BP.
ID	AAV23197	
XX	AAV23197;	
AC	28-UTL-1998	(first entry)
DT	Lactococcus lactis	constitutional promoter Cp3.
XX	Lactococcus lactis;	constitutional promoter; optimise; spacer;
XX	artificial promoter library;	gene expression; de.
XX	Synthetic.	
OS	Lactococcus lactis.	
XX	Key	Location/Qualifiers
FT	promoter	4..60
FT		/tags a
FT		/standard_name= "Constitutional promoter"
XX	W09807846-A1.	
XX	26-FEB-1998.	
XX	25-AUG-1997;	97WO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
XX	(JENS/)	JENSEN P R.
XX	Jensen PR,	Hammer K;
XX	WPI;	1998-179062/16.
PT	New artificial promoter libraries	- containing consensus promoter
PT	sequences and variable spacers,	used to generate promoters for optimising
PT	expression of genes.	
XX	Claim 28;	Page 52; 89pp; English.
XX	This is a Lactococcus lactis	constitutional promoter sequence used in the
XX	construction of an artificial	promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-25

Perfect score: 60

Sequence: 1 CATCTGTAGTATTACTCTG.....TATTAATAGAGAGTACTGT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 2422767955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006956	BD006956 Artificia
2	19	31.7	131000	AC016412	AC016412 Homo sapi
3	19	31.7	175915	AC021235	AC021235 Homo sapi
4	18	30.0	58	BD006943	BD006943 Artificia
5	18	30.0	76329	AB007648	AB007648 Arabidops
6	18	30.0	101261	AC010628	AC010628 Homo sapi
7	18	30.0	127326	AC010682	AC010682 Homo sapi
8	18	30.0	188165	CNS00003	AL049781 Human chr
9	18	30.0	217386	AC106144	AC106144 Rattus no
10	18	30.0	234848	AC109367	AC109367 Rattus no
11	18	30.0	241591	AC131483	AC131483 Rattus no
12	18	30.0	241726	AC131440	AC131440 Rattus no
13	18	30.0	249251	AC134585	AC134585 Mus muscu
14	18	30.0	252458	AC094729	AC094729 Rattus no
15	17	28.3	59	BD006968	BD006968 Artificia
16	17	28.3	60	BD006941	BD006941 Artificia
17	17	28.3	60	BD006964	BD006964 Artificia
18	17	28.3	60	BD006970	BD006970 Artificia
19	17	28.3	60	BD006971	BD006971 Artificia

20	17	28.3	60	BD006973	BD006973 Artificia
21	17	28.3	5118	AY303939	AY303939 Pichia an
22	17	28.3	30254	AL669934	AL669934 Human DNA
23	17	28.3	36184	EX942823	EX942823 Homo sapi
24	17	28.3	55261	AL591468	AL591468 Human DNA
25	17	28.3	58689	AC136343	AC136343 Homo sapi
26	17	28.3	73449	CR556709	CR556709 Zebrafish
27	17	28.3	100660	CO870467	CO870467 Sequence
28	17	28.3	108793	AL137839	AL137839 Human DNA
29	17	28.3	112595	AC020939	AC020939 Homo sapi
30	17	28.3	131575	AC146171	AC146171 Pan trogl
31	17	28.3	135800	AF523316	AF523316 Canis fam
32	17	28.3	137269	AL139137	AL139137 Human DNA
33	17	28.3	151102	AF282130	AF282130 Melalegri
34	17	28.3	154963	AC008911	AC008911 Homo sapi
35	17	28.3	156630	AC146025	AC146025 Pan trogl
36	17	28.3	159160	AF291866	AF291866 Melalegri
37	17	28.3	159610	AC012137	AC012137 Homo sapi
38	17	28.3	165675	AC022163	AC022163 Homo sapi
39	17	28.3	166155	AC093129	AC093129 Papio anu
40	17	28.3	167862	AC011966	AC011966 Homo sapi
41	17	28.3	171783	AC146321	AC146321 Papio anu
42	17	28.3	174174	EX004817	EX004817 Zebrafish
43	17	28.3	176577	AL157831	AL157831 Human DNA
44	17	28.3	183238	AC053525	AC053525 Homo sapi
45	17	28.3	183846	AC073850	AC073850 Homo sapi

ALIGNMENTS

RESULT 1
BD006956
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006956
VERSION
BD006956.1 GI:18635327
KEYWORDS
JP 2001503249-A/25.
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Lactococcus lactis
PN JP 2001503249-A/25
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0866/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
FH Topology: Linear;
FT Key promoter
Location/Qualifiers
(4). (60).

FEATURES
SOURCE
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATCTGTAGTATTACTCTGACACAGGTTAGTGTGATATATAGAGAGTACTGT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 / Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-25

Perfect score: 60
Sequence: 1 CATCTGAGTTATCTTG.....TATATAGAGAGTCTGT 60

Scoring table: OLIGO_NUC

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6847908

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	392	8	AQ304789 HS 2005 B
2	18	30.0	464	8	BH748487 SATK 0460
3	18	30.0	535	8	AQ644596 RPTC193-DP
4	18	30.0	719	7	CO389779 AGNCCOURT
5	17	28.3	439	6	CD195972 MS1-0091T
6	17	28.3	548	8	AQ708599 HS 5362 A
7	17	28.3	572	8	AQ021951 CIT-HSP-2
8	17	28.3	628	9	CL385047 RPTC144 33
9	17	28.3	705	2	BE577417 L48-2257T
10	17	28.3	721	9	CR839711 GROAA8BB
11	17	28.3	739	8	BZ062012 1K67B07
12	17	28.3	745	8	BZ061945 1K67B07
13	17	28.3	779	9	AG314075 Mus muscu
14	17	28.3	789	9	CL390793 ZMMBB019
15	17	28.3	1036	2	BF685539 602142277
16	16	26.7	276	4	BM277425 TM ad 40A
17	16	26.7	283	2	AM885098 OVA-OT006
18	16	26.7	286	2	AM885019 OVA-OT006
19	16	26.7	288	2	BM385117 BM385117
20	16	26.7	307	4	BM385140 NISC m10
21	16	26.7	316	4	BM383196 UT-R-DSO
22	16	26.7	334	1	AV765566 AV765566
23	16	26.7	335	8	BH334308 CH230-260
24	16	26.7	336	5	BP078576 BP078576

c	25	16	26.7	360	7	D75498	D75498 CELK105D1F
c	26	16	26.7	367	8	AZ602750	AZ602750 IM0421J08
c	27	16	26.7	375	5	BY494143	BY494143 BY494143
c	28	16	26.7	380	6	CL3483	CL3483 CL3483 Yuj1
c	29	16	26.7	381	5	BP072143	BP072143 BP072143
c	30	16	26.7	388	1	A1232299	A1232299 EST228987
c	31	16	26.7	389	8	AQ710557	AQ710557 HS 5345 A
c	32	16	26.7	398	5	BP638705	BP638705 BP638705
c	33	16	26.7	418	4	BM418766	BM418766 R004D11 O
c	34	16	26.7	428	5	BP054087	BP054087 BP054087
c	35	16	26.7	429	8	AZ160972	AZ160972 SP 0067 B
c	36	16	26.7	438	5	BP049722	BP049722 BP049722
c	37	16	26.7	441	7	CF255297	CF255297 mdm126 e
c	38	16	26.7	446	1	AA892975	AA892975 EST196778
c	39	16	26.7	449	5	BY386797	BY386797 BY386797
c	40	16	26.7	450	7	CK607709	CK607709 Controlb
c	41	16	26.7	454	5	BP059177	BP059177 BP059177
c	42	16	26.7	459	7	CK610424	CK610424 IRND E13
c	43	16	26.7	460	2	BF408442	BF408442 UT-R-BJ2-
c	44	16	26.7	463	1	AA495704	AA495704 pat PK000
c	45	16	26.7	465	7	CK611066	CK611066 IFN1_D11

ALIGNMENTS

RESULT 1
AQ304789/c 392 bp DNA linear GSS 16-DEC-1998
LOCUS HS 2005 B2 B12 MR CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=2005 Col=24 Row=D, genomic survey
sequence.

ACCESSION AQ304789
VERSION AQ304789.1 GI:4024575
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
1 (bases 1 to 392)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,D., Young,D., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

TITLE
JOURNAL High Throughput Sequencing Center
MEDLINE Contact: Mahairas GG, Wallace JC, Hood L
PUBMED University of Washington
10449764 401 Queen Anne Avenue North, Seattle, WA 98109, USA
99380589 Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

COMMENT
Sequence Tagged Connector
Plate: 2005 row: D column: 24
Class: BAC ends
High quality sequence stop: 392.
Location/Qualifiers
1..392

FEATURES

source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2005 Col=24 Row=D"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"

ORIGIN

Query Match 31.7%; Score 19; DB 8; Length 392;
Best Local Similarity 100.0%; Pred. No. 3;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198,929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-26

Perfect score: 60
Sequence: 1 CATGACAGAGTTTATCTTG.....TATAATAGTTGACTGTT 60

Scoring table:

OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N Geneseq_16Dec04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23198 Lactococc
2	18	30.0	60	2	AAV23177 Lactococc
3	18	30.0	60	2	AAV23189 Lactococc
4	18	30.0	60	2	AAV23204 Lactococc
5	17	28.3	60	2	AAV23191 Lactococc
6	17	28.3	729	9	ADB06849 Altiococc
7	17	28.3	729	9	ADB06851 Altiococc
8	17	28.3	945	5	AA542456 Human CDN
9	17	28.3	945	6	ABK68647 Human DNA
10	17	28.3	945	6	ABK37742 DNA encod
11	17	28.3	945	12	ADG83501 Human Olf
12	17	28.3	947	10	ADG88062 Human GPC
13	17	28.3	947	10	ADG79353 Human GPC
14	17	28.3	1308	9	ADB10947 Altiococc
15	17	28.3	1308	9	ADB12632 Altiococc
16	17	28.3	1308	9	ADB10945 Altiococc
17	17	28.3	1308	9	ADB06373 Altiococc
18	17	28.3	1308	9	ADB09674 Altiococc
19	17	28.3	1308	9	ADB06623 Altiococc
20	17	28.3	1308	9	ADB07657 Altiococc

21	17	28.3	1308	9	ADB06371 Altiococc
22	17	28.3	1308	9	ADB06625 Altiococc
23	17	28.3	1308	9	ADB06929 Altiococc
24	17	28.3	1308	9	ADB06917 Altiococc
25	17	28.3	1308	9	ADB06931 Altiococc
26	17	28.3	1308	9	ADB06919 Altiococc
27	17	28.3	1308	9	ADB12630 Altiococc
28	17	28.3	1308	9	ADB09682 Altiococc
29	17	28.3	1330	10	ADC86242 Human GPC
30	17	28.3	1331	10	ABX17876 CDNA enco
31	17	28.3	13573	6	ABL33868 Human DNA
32	17	28.3	25950	4	AA531518 Human DNA
33	17	28.3	25950	6	AB066842 Human pol
34	17	28.3	25950	10	ADC11129 Human DNA
35	17	28.3	91823	10	ADL13497 Osteocarc
36	17	28.3	110000	9	ADB12064_00 Continuation (2 of
37	17	28.3	110000	9	ADB12064_01 Continuation (2 of
38	17	28.3	110000	9	ADB12064_02 Continuation (3 of
39	17	28.3	110000	9	ADB12064_03 Continuation (4 of
40	17	28.3	110000	9	ADB12064_10 Continuation (11 o
41	17	28.3	110000	9	ADB12064_10 Continuation (17 o
42	17	28.3	110000	9	ADB12064_16 Continuation (17 o
43	17	28.3	19686	11	ACN44170 Human gen
44	17	28.3	19686	11	ACN44170 Human gen
45	17	28.3	198161	6	ABK83564 Human CDN

ALIGNMENTS

RESULT 1	AAV23198	standard; DNA; 60 BP.
ID	AAV23198	standard; DNA; 60 BP.
XX	AAV23198;	
DT	28-JUL-1998	(first entry)
XX	Lactococcus lactis	constitutional promoter Cp30.
DE	Lactococcus lactis	constitutional promoter; optimise; spacer;
XX	Lactococcus lactis	constitutional promoter; optimise; spacer;
KW	artificial promoter library; gene expression; ds.	
XX	Synthetic.	
OS	Lactococcus lactis.	
XX	Lactococcus lactis.	
XX	Key	Location/Qualifiers
FT	promoter	4..60
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
XX	WO9807846-A1.	
XX	26-FEB-1998.	
XX	25-AUG-1997;	97WO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
XX	(JENS/)	JENSEN P R.
XX	Jensen PR,	Hammer K;
XX	WPI, 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX	Claim 28; Page 53; 89pp; English.	
XX	This is a Lactococcus lactis constitutional promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-26

Perfect score: 60
Sequence: 1 CATGACAGAGTTATTCTTG.....TATATAGTTGACTGTT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 segs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBml:.*
1: gb_ba:.*
2: gb_ncg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_str:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	6	BD006957 Artificialia
2	20	33.3	144165	9	AC0002127 Homo sapi
3	20	33.3	169542	9	AC146002 Pan trogl
4	19	31.7	106979	9	AC083804 Homo sapi
5	18	30.0	60	6	BD006936 Artificialia
6	18	30.0	60	6	BD006948 Artificialia
7	18	30.0	60	6	BD006963 Artificialia
8	18	30.0	960	10	AY073514 Mus muscu
9	18	30.0	960	10	AY318240 Mus muscu
10	18	30.0	110000	2	AC101954_1 Continuation (2 of
11	18	30.0	110000	2	AC151279_3 Continuation (4 of
12	18	30.0	119730	9	AC117472 Homo sapi
13	18	30.0	136161	8	AC145330 Medicago
14	18	30.0	151923	10	AL845439 Mouse DNA
15	18	30.0	163863	2	AC102006 Mus muscu
16	18	30.0	169841	9	AL607043 Human DNA
17	18	30.0	170040	5	BX005386 Zebrafish
18	18	30.0	178879	2	AC124052 Mus muscu
19	18	30.0	183632	10	AL929084 Mouse DNA

c	20	18	30.0	188485	2	AC027454 Homo sapi
	21	18	30.0	203354	2	AC107817 Mus muscu
	22	18	30.0	206168	10	AL672307 Mouse DNA
	23	18	30.0	219711	10	AC101969 Mus muscu
	24	18	30.0	219817	2	AC099196 Ratius no
	25	18	30.0	223545	2	AC100545 Mus muscu
	26	18	30.0	251806	2	AC135284 Ratius no
	27	18	30.0	254540	2	AC119383 Ratius no
	28	18	30.0	279974	2	AC112865 Ratius no
	29	18	30.0	287087	2	AC096170 Ratius no
	30	18	30.0	289082	2	AC118813 Ratius no
	31	18	30.0	295566	2	AC122608 Ratius no
	32	18	30.0	318620	2	AC110910 Ratius no
	33	17	28.3	60	6	BD006950 Artificialia
	34	17	28.3	884	10	AY074029 Mus muscu
	35	17	28.3	927	10	AY073836 Mus muscu
	36	17	28.3	936	10	AY318205 Mus muscu
	37	17	28.3	945	6	AX448827 Sequence
	38	17	28.3	946	6	AX702708 Sequence
	39	17	28.3	1330	6	AX646503 Sequence
	40	17	28.3	1330	9	AB065603 Homo sapi
	41	17	28.3	1846	9	AK093511 Homo sapi
	42	17	28.3	2708	9	L80211 Homo sapien
	43	17	28.3	2900	8	AK119463 Oryza sat
	44	17	28.3	8661	8	SPAC631 S.pombe c
	45	17	28.3	13573	6	AX346770 Sequence

ALIGNMENTS

RESULT 1	BD006957	Artificial promoter libraries for selected organisms and promoters derived from such libraries.	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006957					
DEFINITION	BD006957	Artificial promoter libraries for selected organisms and promoters derived from such libraries.				
ACCESSION	BD006957.1	GI:18635328				
VERSION	JP 2001503249-A/26					
KEYWORDS	Lactococcus lactis					
SOURCE	Lactococcus lactis					
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.					
REFERENCE	1 (bases 1 to 60)					
AUTHORS	Hammer, K. and Janssen, P.R.					
TITLE	Artificial promoter libraries for selected organisms and promoters derived from such libraries					
JOURNAL	Patent: JP 2001503249-A 26 13-MAR-2001;					
COMMENT	PETER RUDAR JANSSEN					
OS	Lactococcus lactis					
PN	JP 2001503249-A/26					
PD	13-MAR-2001					
PF	25-AUG-1997 JP 1998510287					
PR	23-AUG-1996 DK 0886/96					
PI	KALIN HAMMER, PETER RUDAR JANSSEN					
PC	CI2N15/09, CI2N15/00					
CC	Strandedness: Double;					
CC	Topology: linear;					
CC	Key					
FT	promoter (4) ..(60).					
FEATURES	Location/Qualifiers					
source	1..60					
	/organism="Lactococcus lactis"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:1358"					
ORIGIN						
Query Match	100.0%; Score 60; DB 6; Length 60;					
Best Local Similarity	100.0%; Pred. No. 1e-21;					
Matches	60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
1	CATGACAGAGTTATTCTTGACAGATGGGTTACTTGTATATAGTTGACTGTT 60					

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds

(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-26

Perfect score: 60
Sequence: 1 CATGACAGAGTTTATCTGTG.....TATATAGTTAGACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	33.3	542	5	BU071968 im52f01.y
2	20	33.3	786	7	CN762383 ID0AA3DH
3	18	30.0	424	8	A2290770 RPCI-23-1
4	18	30.0	475	5	BQ161553 WHE1455 B
5	18	30.0	518	6	CB470284 sm20 A0B
6	18	30.0	601	4	B3133250 B3133250
7	18	30.0	613	6	CB076546 h146f08.g
8	18	30.0	648	9	CR236800 Forward B
9	18	30.0	664	8	A2018318 RPCI-23-2
10	18	30.0	671	6	CA159481 SCJFR23C0
11	18	30.0	674	6	CA159393 SCJFR23C0
12	18	30.0	854	8	B2137209 CH230-386
13	18	30.0	860	7	CK091218 F061P27.3
14	17	28.3	310	7	CN941951 010917AVB
15	17	28.3	346	2	BB346297 BB346297
16	17	28.3	368	7	CN634247 ta106D03
17	17	28.3	368	7	CO754708 Mdfrc1051
18	17	28.3	372	7	CO576687 Mdfrc1042
19	17	28.3	378	7	CO576708 Mdfrc1042
20	17	28.3	380	7	CN774458 ta681e10
21	17	28.3	383	7	CN876118 020813AAR
22	17	28.3	383	7	CN876826 020814AAR
23	17	28.3	386	6	CA913870 PC600790X
24	17	28.3	390	7	CO541321 Mdfrc1039

25	17	28.3	394	7	CN994676	CN994676 Mdfw2054m
26	17	28.3	404	6	CD465864	CD465864 leucon1.8
27	17	28.3	408	7	CN874933	CN874933 010130AAR
28	17	28.3	415	7	CN634449	CN634449 ta106D03
29	17	28.3	417	8	AQ173773	AQ173773 HS_3194_A
30	17	28.3	427	8	AQ220278	AQ220278 HS_3250_B
31	17	28.3	450	8	AQ225650	AQ225650 HS_2003_B
32	17	28.3	525	7	CV161937	CV161937 CS_g11_79
33	17	28.3	535	7	CN635151	CN635151 11B01_54
34	17	28.3	545	6	CA872326	CA872326 K0917A04
35	17	28.3	551	7	CF132821	CF132821 4-03-D-Ch
36	17	28.3	555	6	CB173910	CB173910 OR_2019A0
37	17	28.3	563	8	AO855238	AO855238 CPG1863A
38	17	28.3	569	7	CF808235	CF808235 p8H033xE
39	17	28.3	569	7	CK541566	CK541566 tswhb0.00
40	17	28.3	576	7	CN557437	CN557437 ta49h10
41	17	28.3	580	9	CPA562616	CPA562616 Cryptospo
42	17	28.3	581	7	CF426773	CF426773 Lr_LC1ED
43	17	28.3	599	7	CN627191	CN627191 ta689e09
44	17	28.3	601	9	CL718765	CL718765 OR_B8A004
45	17	28.3	602	8	AQ509292	AQ509292 nbxb0094A

ALIGNMENTS

RESULT 1
BU071968 542 bp mRNA linear EST 27-AUG-2002
im52f01.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6038904 5',
mRNA sequence.
BU071968
BU071968.1 GI:22513157

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2000)
Endocrine Pancreas Consortium
Other ESTs: im52f01.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@chp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seg primer: -40RP from Gibco
High quality sequence stop: 481.
Location/Qualifiers
1..542
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6038904"
/feature_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-27

Sequence: 1 .CATACGGAGTTATCTTG.....TATAATACTTAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 200000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04: *

```

1:  gensesegm1980bs: *
2:  gensesegm1998bs: *
3:  gensesegm2000bs: *
4:  gensesegm2001bs: *
5:  gensesegm2001bs: *
6:  gensesegm2002bs: *
7:  gensesegm2002bs: *
8:  gensesegm2003bs: *
9:  gensesegm2003bs: *
10: gensesegm2003cs: *
11: gensesegm2003ds: *
12: gensesegm2004bs: *
13: gensesegm2004bs: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV23199	AAV23199 Lactococcus
2	20	33.3	185371	6	ABT10718	Abt10718 Human bre
3	19	31.7	59	2	AAV23210	AAV23210 Lactococcus
4	19	31.7	60	2	AAV23200	AAV23200 Lactococcus
5	18	30.0	60	2	AAV23211	AAV23211 Lactococcus
6	17	28.3	60	2	AAV23177	AAV23177 Lactococcus
7	17	28.3	60	2	AAV23191	AAV23191 Lactococcus
8	17	28.3	58	2	AAV23186	AAV23182 Lactococcus
9	16	26.7	59	2	AAV23182	AAV23186 Lactococcus
10	16	26.7	59	2	AAV23190	AAV23190 Lactococcus
11	16	26.7	59	2	AAV23185	AAV23185 Lactococcus
12	16	26.7	59	2	AAV23193	AAV23193 Lactococcus
13	16	26.7	59	2	AAV23206	AAV23206 Lactococcus
14	16	26.7	60	2	AAV23194	AAV23194 Lactococcus
15	16	26.7	60	2	AAV23188	AAV23188 Lactococcus
16	16	26.7	60	2	AAV23189	AAV23189 Lactococcus
17	16	26.7	60	2	AAV23198	AAV23198 Lactococcus
18	16	26.7	60	2	AAV23212	AAV23198 Lactococcus
19	16	26.7	60	6	ABK98622	Abk98622 Lambda CP
20	16	26.7	60	9	ACD13873	Acid13873 L. lactis

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-27

Perfect score: 60

Sequence: 1. CATACGGAGTTATCTCTG.....TATAATTAAGTACTGTT 60

Scoring table: OLIGO_NTC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_dr:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006958	BD006958 Artificia
2	20	33.3	185371	6 HS625H18	AL022726 Human DNA
3	19	31.7	59	6 BD006969	BD006969 Artificia
4	19	31.7	60	6 BD006959	BD006959 Artificia
5	19	31.7	177104	10 AC091237	AC091237 Mus muscu
6	19	31.7	184810	10 AC090869	AC090869 Mus muscu
7	19	31.7	260036	2 AC130079	AC130079 Rattus no
8	18	30.0	60	6 BD006970	BD006970 Artificia
9	18	30.0	2457	5 BC082373	BC082373 Xenopus l
10	18	30.0	122633	10 BX005236	BX005236 Mouse DNA
11	18	30.0	146927	2 AC020144	AC020144 Drosophila
12	18	30.0	157722	3 AC092248	AC092248 Drosophila
13	18	30.0	16978	9 AC092248	AC092248 Homo sapi
14	18	30.0	174974	3 AC092228	AC092228 Drosophila
15	18	30.0	234075	2 AC127149	AC127149 Rattus no
16	18	30.0	226144	3 AB003643	AB003643 Drosophila
17	18	30.0	240584	2 AC116063	AC116063 Rattus no
18	18	30.0	255239	2 AC109676	AC109676 Rattus no
19	18	30.0	305502	3 DROSADH03	AE003409 Drosophila

20	17	28.3	60	6 BD006936	BD006936 Artificia
21	17	28.3	60	6 BD006941	BD006941 Artificia
22	17	28.3	60	6 BD006950	BD006950 Artificia
23	17	28.3	1470	8 AK119468	AK119468 Oryza sat
24	17	28.3	3378	8 AK095955	AK095955 Oryza sat
25	17	28.3	4209	8 AB178084	AB178084 Daucus ca
26	17	28.3	35688	9 AC073417	AC073417 Homo sapi
27	17	28.3	57730	2 AC014848	AC014848 Drosophila
28	17	28.3	66273	9 AC004892	AC004892 Homo sapi
29	17	28.3	72790	9 AC108930	AC108930 Homo sapi
30	17	28.3	84076	2 AC149797	AC149797 Aedes aeg
31	17	28.3	85624	9 AL139393	AL139393 Human DNA
32	17	28.3	91204	2 BX927126	BX927126 Mus muscu
33	17	28.3	96877	2 AC017776	AC017776 Drosophila
34	17	28.3	107600	2 APL16313	APL16313 Arabidops
35	17	28.3	110000	2 AC128164	AC128164 Rattus no
36	17	28.3	110000	2 AL845451_2	Continuation (3 of
37	17	28.3	110000	2 AB016817_04	Continuation (5 of
38	17	28.3	110440	9 HS293514	HS293514 Human DNA s
39	17	28.3	110441	2 HS23	HS23 Homo sapien
40	17	28.3	114655	9 AC079178	AC079178 Homo sapi
41	17	28.3	129414	8 AC140545	AC140545 Medicago
42	17	28.3	137272	9 AL137062	AL137062 Human DNA
43	17	28.3	138824	8 AC051633	AC051633 Oryza sat
44	17	28.3	139505	9 HSJ591C20	AL118506 Human DNA
45	17	28.3	142114	8 AC073166	AC073166 Oryza sat

ALIGNMENTS

RESULT 1	BD006958	60 bp	DNA	linear	PAT 31-JUN-2002
LOCUS	BD006958	Artificial promoter libraries for selected organisms and promoters			
DEFINITION	BD006958	derived from such libraries.			
ACCESSION	BD006958.1	GI:18635329			
VERSION	JP 2001503249-A/27.				
KEYWORDS	JP 2001503249-A/27.				
SOURCE	Lactococcus lactis				
ORGANISM	Lactococcus lactis				
REFERENCE	1 (bases 1 to 60)				
AUTHORS	Hammer, K. and Janssen, P.R.				
TITLE	Artificial promoter libraries for selected organisms and promoters				
JOURNAL	Patent: JP 2001503249-A 27 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
OS	Lactococcus lactis				
PN	JP 2001503249-A/27				
PD	13-MAR-2001				
PR	25-AUG-1997 JP 1998510287				
PI	23-AUG-1996 DK 0886/96				
PC	KALIN HAMMER, PETER RUDAR JANSSEN				
CC	C12N15/09, C12N15/00				
CC	Strandedness: Double;				
CC	Topology: linear;				
FT	Key				
FEATURES	Location/Qualifiers				
source	1..60				
	/organism="Lactococcus lactis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN	Query Match	100.0%;	Score 60;	DB 6;	Length 60;
	Best Local Similarity	100.0%;	Pred. No. 4.8e-23;		
	Matches	60;	Conservative	0;	Mismatches 0;
					Indels 0;
					Gaps 0;
QY	1 CATACGGAGTTATCTTGACATATGCCGGTGTGTGATATTAAGTACTGTT 60				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-27

Perfect score: 60

Sequence: 1 CATACGAGGATTATCTTCTT.....TATAATTAAGTACTAGTGT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	723	9	CC848135 NDL.112K2
2	18	30.0	321	2	BE491516 db27b07.x
3	18	30.0	358	6	CF029617 QCD12a02.
4	18	30.0	419	5	BP715056 BP715056
5	18	30.0	420	8	AZ660372 IM0538X21
6	18	30.0	432	4	BI315179 dah95a08.
7	18	30.0	436	6	CF030799 QCD26f04.
8	18	30.0	436	6	CF030858 QCD27d07.
9	18	30.0	436	6	CF030875 QCD27f06.
10	18	30.0	436	6	CF031004 QCD29c09.
11	18	30.0	436	6	CF031256 QCD29c02.
12	18	30.0	437	6	CF030997 QCD29c02.
13	18	30.0	438	6	CF029631 QCD12a02.
14	18	30.0	438	6	CF030356 QCD21a04.
15	18	30.0	438	6	CF030598 QCD24b05.
16	18	30.0	456	6	CF058870 QCD24b04.
17	18	30.0	457	6	CF030597 QCD24b04.
18	18	30.0	458	1	AI855241 603010G10
19	18	30.0	460	6	CF030874 QCD27f05.
20	18	30.0	467	6	CD988414 QCD27f05.
21	18	30.0	470	5	BP677992 BP677992
22	18	30.0	473	6	CF030752 QCD26a08.
23	18	30.0	475	5	BQ161553 WHB145_B
24	18	30.0	479	6	CF030599 QCD24b08.

c 25	18	30.0	484	6	CF031602 QCD7b01.Y
c 26	18	30.0	485	6	CF030870 QCD27e12.
c 27	18	30.0	486	6	CF031026 QCD29f04.
c 28	18	30.0	486	6	CF031591 QCD7a02.Y
c 29	18	30.0	510	6	CF031804 QCD9g01.Y
c 30	18	30.0	515	2	BE761596 947003B06
c 31	18	30.0	540	6	CF029997 QCD17d03.
c 32	18	30.0	541	5	BM953850 sam67a04.
c 33	18	30.0	543	6	CF030767 QCD26c01.
c 34	18	30.0	543	6	CF031116 QCD2h10.Y
c 35	18	30.0	544	6	CF030447 QCD22b05.
c 36	18	30.0	544	6	CF058624 QCD12d12.
c 37	18	30.0	545	6	CF030120 QCD19a04.
c 38	18	30.0	545	6	CF030287 QCD20b01.
c 39	18	30.0	546	6	CF029723 QCD13h09.
c 40	18	30.0	546	6	CF029916 QCD16d02.
c 41	18	30.0	549	6	CD979358 QAG2f06.Y
c 42	18	30.0	550	6	CF029674 QCD13c09.
c 43	18	30.0	550	6	CF030335 QCD20g02.
c 44	18	30.0	551	6	CF030202 QCD1a06.Y
c 45	18	30.0	551	6	CF031317 QCD3b10.Y

ALIGNMENTS

RESULT 1
LOCUS CC848135 723 bp DNA linear GSS 24-JUL-2003
DEFINITION NDL.112K23.T7 Notre Dame Liverpool Aedes aegypti genomic clone
Notredame Liverpool-112K23, genomic survey sequence.
ACCESSION CC848135
VERSION CC848135.1 GI:33196744
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Chiloidea; Aedes; Stegomyia.

REFERENCE 1 (bases 1 to 723)
Loftus, B., Shetty, J., Knudson, D. and Severson, D.
TITLE BAC end sequencing of Aedes aegypti
JOURNAL Unpublished (2003)
COMMENT Other GSSs: NDL.112K23.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.

FEATURES

source 1..723
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="Notredame Liverpool-112K23"
/clone_lib="Notredame Liverpool"
/note="Vector: pGCBAC1; Site_1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 31.7%; Score 19; DB 9; Length 723;
Best Local Similarity 100.0%; Pred.No. 3.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 AGTTATCTTGACATATT 27

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds

(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-28

Perfect score: 60
Sequence: 1 CATGTGAGACTTATCTCTG.....TATATAGGTGAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1980s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	AAV23200	AAV23200 Lactococc
2	23	38.3	60	AAV23213	AAV23213 Lactococc
3	19	31.7	60	AAV23199	AAV23199 Lactococc
4	18	30.0	59	AAV23210	AAV23210 Lactococc
5	17	28.3	59	AAV23185	AAV23185 Lactococc
6	17	28.3	60	AAV23211	AAV23211 Lactococc
7	17	28.3	60	AAV23177	AAV23177 Lactococc
8	17	28.3	60	AAV23191	AAV23191 Lactococc
9	17	28.3	60	AAV23182	AAV23182 Lactococc
10	17	28.3	3888	13	ADT91410
11	17	28.3	27067	10	ADT91410
12	17	28.3	38239	12	ADT91410
13	17	28.3	185371	6	ADT91410
14	17	26.7	58	2	AAV23186
15	16	26.7	58	2	AAV23190
16	16	26.7	59	2	AAV23193
17	16	26.7	59	2	AAV23206
18	16	26.7	60	2	AAV23194
19	16	26.7	60	2	AAV23188
20	16	26.7	60	2	AAV23189

ALIGNMENTS

21	16	26.7	60	2	AAV23198	AAV23198 Lactococc
22	16	26.7	60	2	AAV23212	AAV23212 Lactococc
23	16	26.7	60	6	ABK98622	ABK98622 Lambda CP
24	16	26.7	60	9	ACD13873	ACD13873 L. lactis
25	16	26.7	95	6	ABK98612	ABK98612 Lambda CP
26	16	26.7	95	6	ACD13863	ACD13863 L. lactis
27	16	26.7	265	2	AAQ23880	AAQ23880 PKTH1820
28	16	26.7	265	2	AAQ46205	AAQ46205 Promoter
29	16	26.7	265	2	AAQ31875	AAQ31875 Promoter
30	16	26.7	296	3	AAQ31875	AAQ31875 Human sec
31	16	26.7	460	9	AAQ31875	AAQ31875 Human sec
32	16	26.7	466	9	AAQ31875	AAQ31875 Human sec
33	16	26.7	747	6	ABQ17753	ABQ17753 Oligonuc
34	16	26.7	747	6	ABQ17753	ABQ17753 Oligonuc
35	16	26.7	748	6	ABQ18626	ABQ18626 Oligonuc
36	16	26.7	748	6	ABQ18627	ABQ18627 Oligonuc
37	16	26.7	835	8	ABZ35933	ABZ35933 Human sec
38	16	26.7	841	4	AAQ07905	AAQ07905 Human sec
39	16	26.7	876	5	AAQ84885	AAQ84885 DNA encod
40	16	26.7	1073	4	AAH19197	AAH19197 Human sec
41	16	26.7	1134	6	ABT07897	ABT07897 Human lun
42	16	26.7	1354	8	ABZ18320	ABZ18320 Group III
43	16	26.7	1427	4	AAH19199	AAH19199 Human sec
44	16	26.7	1550	5	AAQ27276	AAQ27276 Mouse atc
45	16	26.7	1550	9	AAQ62223	AAQ62223 Mouse atc

RESULT 1

ID AAV23200 standard; DNA; 60 BP.

AC AAV23200;

DT 28-JUL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cp33.

KW Lactococcus lactis; constitutional promoter; optimise; spacer;

OS Lactococcus lactis.

OS Synthetic.

FT Key Location/Qualifiers

FT promoter 4..60 /tag= a

PN WO9807846-A1.

PD 26-FEB-1998.

PF 25-AUG-1997; 97WO-DK000342.

PR 23-AUG-1996; 96DK-00000886.

PA (JENSEN) JENSEN P R.

PI Jensen PR, Hammer K;

DR WPI; 1998-179062/16.

PT New artificial promoter libraries - containing consensus promoter

PS sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.

CC Claim 28; Page 54; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657b-28

Perfect score: 60

Sequence: 1 CATGTGACGATTATCTCTG.....TATATAGTGAGTACTGTT 60

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl:
1: gb_ba:*
2: gb_breg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006959	BD006959 Artificia
2	23	38.3	60	BD006972	BD006972 Artificia
3	19	31.7	60	BD006958	BD006958 Artificia
4	19	31.7	169978	AC017013	AC017013 Homo sapi
5	19	31.7	172826	AL671917	AL671917 Mouse DNA
6	19	31.7	184147	AC103805	AC103805 Homo sapi
7	19	31.7	203630	BX936444	BX936444 Dario rer
8	19	31.7	233120	AC132712	AC132712 Rattus no
9	19	31.7	260779	CR382368	CR382368 Dario rer
10	19	31.7	272900	AC145760	AC145760 Microcebu
11	18	30.0	36215	BD006969	BD006969 Artificia
12	18	30.0	65476	AY330343	AY330343 Odocoileu
13	18	30.0	84076	AC149797	AC149797 Aedes aeg
14	18	30.0	157267	AC007859	AC007859 Homo sapi
15	18	30.0	164709	CR388063	CR388063 Dario rer
16	18	30.0	165993	AL172362	AL172362 Zebrafish
17	18	30.0	196548	AC120074	AC120074 Rattus no
18	18	30.0	202566	AC022031	AC022031 Homo sapi

20	18	30.0	215731	2	AC136554	AC136554 Rattus no
21	18	30.0	230071	2	AC132721	AC132721 Rattus no
22	18	30.0	230514	2	AC106233	AC106233 Rattus no
23	18	30.0	234163	2	AC115318	AC115318 Rattus no
24	18	30.0	240238	2	AC095867	AC095867 Rattus no
25	18	30.0	254007	2	AC094413	AC094413 Rattus no
26	18	30.0	254483	2	AC097072	AC097072 Rattus no
27	18	30.0	266762	2	AC109853	AC109853 Rattus no
28	18	30.0	266999	2	AC121640	AC121640 Rattus no
29	17	28.3	59	6	BD006944	BD006944 Artificia
30	17	28.3	60	6	BD006936	BD006936 Artificia
31	17	28.3	60	6	BD006941	BD006941 Artificia
32	17	28.3	60	6	BD006950	BD006950 Artificia
33	17	28.3	60	6	BD006970	BD006970 Artificia
34	17	28.3	608	11	BD067407	BD067407 S12B6920
35	17	28.3	1470	8	AK119468	AK119468 Oryza sat
36	17	28.3	3378	8	AK099595	AK099595 Oryza sat
37	17	28.3	16141	8	AP006706	AP006706 Lotus cor
38	17	28.3	31188	3	U67951	U67951 Caenorhabdi
39	17	28.3	34496	3	U41016	U41016 Caenorhabdi
40	17	28.3	72767	2	AC149517	AC149517 Xenopus l
41	17	28.3	83232	8	AP004503	AP004503 Lotus cor
42	17	28.3	85624	9	AL139393	AL139393 Human DNA
43	17	28.3	89925	5	BX323059	BX323059 Zebrafish
44	17	28.3	92347	5	BX001037	BX001037 Zebrafish
45	17	28.3	94570	8	AP006713	AP006713 Lotus cor

ALIGNMENTS

RESULT 1
BD006959

DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.

ACCESSION
BD006959
VERSION
BD006959.1 GI:18635330

KEYWORDS
JP 2001503249-A/28.
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE

1 (bases 1 to 60)
AUTHORS
Hammer, K. and Janssen, P.R.
TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 28 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS
Lactococcus lactis
PN
JP 2001503249-A/28
PD
13-MAR-2001
PF
25-AUG-1997 JP 1998510287
PR
23-AUG-1996 DK 0886/96
PI
KALIN HAMMER, PETER RUDAR JANSSEN
PC
C12N15/09, C12N15/00
CC
Strandness: Double;
CC
Topology: Linear;
FH
Key
FT
promoter
Location/Qualifiers
(4). (60).

FEATURES

source
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0% Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGTGACGATTATCTTACATCAATTAAGTACGATGATATTAAGTACTGTT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-28

Perfect score: 60
Sequence: 1 CATGTGGAGTTATCTCTG.....TATTAATGAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	31.7	445	5	BX282567
2	19	31.7	661	4	B1597545
3	18	30.0	170	4	BG958678
4	18	30.0	307	6	CA276845
5	18	30.0	407	6	CB771598
6	18	30.0	417	8	B2156236
7	18	30.0	424	8	B15727
8	18	30.0	433	7	CV293192
9	18	30.0	475	5	BQ161553
10	18	30.0	545	8	AO720075
11	18	30.0	582	8	BH368974
12	18	30.0	590	8	BH375740
13	18	30.0	615	7	CF575488
14	18	30.0	625	8	BH333458
15	18	30.0	635	8	B2139446
16	18	30.0	662	7	CV293829
17	18	30.0	691	8	BH9247283
18	18	30.0	718	9	BX208894
19	18	30.0	744	6	CA217560
20	18	30.0	750	8	BZ058041
21	18	30.0	790	8	CC863504
22	18	30.0	817	9	CG768533
23	18	30.0	829	8	BZ209454
24	18	30.0	841	6	CA065609

25	18	30.0	883	7	CK311400	CK311400
26	17	28.3	226	2	BB497145	BB497145
27	17	28.3	239	8	AZ711212	RPDI-24-1
28	17	28.3	254	9	CE373086	t1gr-98s-
29	17	28.3	257	9	CG604096	OST279459
30	17	28.3	321	2	BB270610	BB270610
31	17	28.3	391	9	AG244489	AG244489
32	17	28.3	394	5	BY371615	BY371615
33	17	28.3	405	5	BY371805	BY371805
34	17	28.3	446	8	AO138798	HS 3071 B
35	17	28.3	489	6	CA520010	CA520010
36	17	28.3	504	9	CR314320	CR314320
37	17	28.3	514	6	CA524826	KS12043H0
38	17	28.3	516	9	CE277651	CE277651
39	17	28.3	547	5	BW550106	BW550106
40	17	28.3	549	7	CV293841	EST882218
41	17	28.3	567	9	CE019656	CE019656
42	17	28.3	601	4	Bj1313250	Bj1313250
43	17	28.3	621	6	CA847459	CA847459
44	17	28.3	628	7	CF790646	CF790646
45	17	28.3	633	7	CK753818	eca01-13c

ALIGNMENTS

RESULT 1
BX282567
LOCUS
DEFINITION
IMAGE:5276908, mRNA sequence.
BX282567.1 GI:28615268
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 445)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, B., Peters, M.,
Radloff, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGP998E0511699.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No. 972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?responsefileNo=972
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heinrichweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACACAGGAACACTATGAC.

FEATURES

source
1..445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGP998E0511699 ; IMAGE:5276908"
/feature_type="hippocampus"
/lab_host="RDH10B"
/clone_lib="NTH_MGC_95"
/note="Organ: Brain; Vector: Bluescript (modified
bluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gcgag); Oligo-dr primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-29

Perfect score: 60

Sequence: 1 CATCGCGAGTTATCTTC.....TATAATACACAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	60	100.0	AAV23201	AAV23201 Lactococ
2	19	31.7	AD895005	AD895005 Human the
3	19	31.7	AD895167	AD895167 Human the
4	17	28.3	AAV23192	AAV23192 Lactococ
5	17	28.3	AD059434	AD059434 Human can
6	16	26.7	AAV23202	AAV23202 Lactococ
7	16	26.7	ABN70992	ABN70992 Streptoco
8	16	26.7	ABN67348	ABN67348 Streptoco
9	16	26.7	AAH16145	AAH16145 Human CDN
10	16	26.7	AA159090	AA159090 Human pol
11	16	26.7	ADQ99313	ADQ99313 DNA encod
12	16	26.7	AD849073	AD849073 Novel hum
13	16	26.7	AD85186	AD85186 Human tum
14	16	26.7	ACN39013	ACN39013 Breast ca
15	16	26.7	ABR25849	ABR25849 Human pol
16	16	26.7	ABR11070	ABR11070 Drosophi
17	16	26.7	AA160876	AA160876 Human pol
18	16	26.7	AD862327	AD862327 Human CDN
19	16	26.7	ADR07912	ADR07912 Full leng
20	16	26.7	ADR07827	ADR07827 Full leng

21	16	26.7	4810	12	ADJ95416	ADJ95416 Human noo
22	16	26.7	5796	4	AB118716	AB118716 Drosophi
23	16	26.7	5897	13	ADR07823	ADR07823 Full leng
24	16	26.7	10594	4	AAK65079	AAK65079 Human imm
25	16	26.7	20165	4	AD197979	AD197979 Human ST
26	16	26.7	110000	6	ABN71527	ABN71527 Human ST
27	16	26.7	110000	13	ABD32921_6	ABD32921_6 Continuation (12 o
28	15	25.0	60	2	AAV23204	AAV23204 Lactococ
29	15	25.0	84	4	AB843403	AB843403 Human liv
30	15	25.0	131	7	AD868868	AD868868 Corn seed
31	15	25.0	155	4	AA155981	AA155981 Probe #24
32	15	25.0	155	4	AAK23972	AAK23972 Human liv
33	15	25.0	155	4	AB849685	AB849685 Human liv
34	15	25.0	188	3	AA183356	AA183356 Human sec
35	15	25.0	295	8	ABX52438	ABX52438 Bovine ES
36	15	25.0	300	2	AA109214	AA109214 Virulence
37	15	25.0	322	8	AB217722	AB217722 S2. subtra
38	15	25.0	343	6	ABK29810	ABK29810 Colon ade
39	15	25.0	343	6	ABK29495	ABK29495 Colon ade
40	15	25.0	381	4	AB830422	AB830422 Human liv
41	15	25.0	393	6	ABN24358	ABN24358 Human ORF
42	15	25.0	433	6	ABA01728	ABA01728 Apple chl
43	15	25.0	459	3	AAF09147	AAF09147 Fusarium
44	15	25.0	463	6	ABO58707	ABO58707 Human col
45	15	25.0	476	4	AA192091	AA192091 Human pol

ALIGNMENTS

RESULT 1

AAV23201 standard; DNA; 60 BP.

AAV23201;

28-JUL-1998 (first entry)

Lactococcus lactis constitutonal promoter Cp34.

Lactococcus lactis; constitutonal promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Synthetic.

Lactococcus lactis.

Key Location/Qualifiers

promoter 4..60

26-FEB-1998.

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENS/) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28, Page 54; 89pp; English.

This is a Lactococcus lactis constitutonal promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-29

Perfect score: 60
Sequence: 1 CATGGCGAGTTATCTCTC.....TATATACACAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hng:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_yt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	BD006960	BD006960 Artificial
2	19	31.7	17144	1 AEO01473	AEO01473 Helicobac
3	19	31.7	95848	5 BX510930	BX510930 Zebrafish
4	19	31.7	159512	2 AC133552	AC133552 Homo sapi
5	19	31.7	161970	2 AC131648	AC131648 Homo sapi
6	19	31.7	166297	2 AC011271	AC011271 Homo sapi
7	19	31.7	170517	2 AC010545	AC010545 Homo sapi
8	19	31.7	181720	9 AC099535	AC099535 Homo sapi
9	19	31.7	192296	2 CR753895	CR753895 Dario rer
10	19	31.7	329000	3 TB8CHRI13	TB8CHRI13 Trypanoso
11	18	30.0	368	14 AF191382	AF191382 HIV-1 iso
12	18	30.0	663	14 AY076045	AY076045 HIV-1 iso
13	18	30.0	1953	5 XU043524	XU043524 Xenopus lae
14	18	30.0	4256	5 XU043523	XU043523 Xenopus lae
15	18	30.0	84274	5 BX470141	BX470141 Zebrafish
16	18	30.0	108799	9 AC090989	AC090989 Homo sapi
17	18	30.0	152520	2 AC148269	AC148269 Callithrix
18	18	30.0	163287	10 AC122290	AC122290 Mus muscu
19	18	30.0	175016	5 BX005065	BX005065 Zebrafish

20	18	30.0	176148	9 AC010798	AC010798 Homo sapi
21	18	30.0	178825	2 AP001203	AP001203 Homo sapi
22	18	30.0	185953	5 BX511144	BX511144 Zebrafish
23	18	30.0	187938	2 AC117973	AC117973 Rattus no
24	18	30.0	189592	2 AC148267	AC148267 Callithrix
25	18	30.0	252420	3 AE014841	AE014841 Plasmodiu
26	18	30.0	268049	2 AC129466	AC129466 Rattus no
27	18	30.0	325612	2 AC095761	AC095761 Rattus no
28	17	28.3	60	6 BD006951	BD006951 Artificial
29	17	28.3	488	14 AY543558	AY543558 HIV-1 iso
30	17	28.3	503	14 AY543556	AY543556 HIV-1 iso
31	17	28.3	599	14 H1M429865	H1M429865 Human imm
32	17	28.3	666	5 CR389790	CR389790 Gallus ga
33	17	28.3	2550	8 AY612790	AY612790 Cruchinima
34	17	28.3	8793	14 AF484518	AF484518 HIV-1 iso
35	17	28.3	46373	9 AC019110	AC019110 Homo sapi
36	17	28.3	50129	2 AC117993	AC117993 Clona sav
37	17	28.3	52811	2 AC099978	AC099978 Mus muscu
38	17	28.3	54618	9 AC104800	AC104800 Homo sapi
39	17	28.3	103610	9 HSD240B8	AL048544 Human DNA
40	17	28.3	106123	9 AL158015	AL158015 Human DNA
41	17	28.3	108821	9 AL161791	AL161791 Human DNA
42	17	28.3	110000	2 AC110125	Continuation (2 of
43	17	28.3	110000	2 LMFCHR25_3	Continuation (4 of
44	17	28.3	118150	9 H5J800C24	AL121593 Human DNA
45	17	28.3	119956	2 AC015789	AC015789 Homo sapi

ALIGNMENTS

RESULT 1
BD006960
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006960
ACCESSION
BD006960.1 GI:18635331
VERSION
JP 2001503249-A/29.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
1 (bases 1 to 60)
AUTHORS
Hammer, R. and Janssen, P.R.
TITLES
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 29 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT
OS Lactococcus lactis
PN JP 2001503249-A/29
PD 13-MAR-2001
PP 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter
Location/Qualifiers
(4) ..(60).

FEATURES
Source
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN
Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-17;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGGCGAGTTATCTCTCACACACCGAGAGCTGTGTATATACACAGTACTGTT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-29

Perfect score: 60
Sequence: 1 CATCGCGAAGTTTATCTTC.....TATAATCAACAGACTCTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gse1:
9: gb_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	33.3	1012	9	CL055958 CH216-81P
C 2	19	31.7	370	2	BE165847 CMO-HT048
C 3	19	31.7	370	2	BE165847 CMO-HT048
C 4	19	31.7	565	9	BE155595 Danio rer
C 5	19	31.7	629	8	AO940350 Shared D
C 6	19	31.7	920	5	BU794913 SUR2D1G07
C 7	18	30.0	424	4	BS613928 dat37b12.
C 8	18	30.0	487	8	B2752666 PUDCH89TB
C 9	18	30.0	524	9	CL003689 gba1 VM5
C 10	18	30.0	536	8	AZ876844 2M0191P15
C 11	18	30.0	593	9	CE325174 t1gr-g88-
C 12	18	30.0	657	8	AZ209236 SP_0105_A
C 13	18	30.0	701	9	CE719135 t1gr-g88-
C 14	18	30.0	712	9	CE719135 t1gr-g88-
C 15	18	30.0	751	5	CC832859 ZMMBD018
C 16	18	30.0	883	7	BX850994 BX850994
C 17	18	30.0	886	6	CK798944 AGENCOURT
C 18	18	30.0	891	6	CD329761 AGENCOURT
C 19	18	30.0	1009	6	CD255500 AGENCOURT
C 20	18	30.0	1054	2	AL342199 Tetradon
C 21	18	30.0	1064	2	BE784583 BM906718
C 22	18	30.0	1064	2	BE784583 BM906718
C 23	18	30.0	1877	5	BM906718 AGENCOURT
C 24	18	30.0	3861	3	BC046944 Xenopus 1
C 25	17	28.3	370	9	AL946714 Arabidops

C 25	17	28.3	402	8	AQ219870 HS_3228_B
C 26	17	28.3	515	8	AQ166888 HS_3166_A
C 27	17	28.3	524	8	AZ002397 RPT-23-3
C 28	17	28.3	538	1	AL925454 AL925454
C 29	17	28.3	540	6	CD730390 4038783_1
C 30	17	28.3	550	8	AO451995 HS_5186_B
C 31	17	28.3	556	8	AQ223568 HS_2011_B
C 32	17	28.3	561	8	AQ2154350 HS_3006_A
C 33	17	28.3	573	8	B2304472 K02967.P1
C 34	17	28.3	603	9	PT017E05R Parametiu
C 35	17	28.3	611	7	CV097969 FMU USDA
C 36	17	28.3	619	5	BO135146 INTT_3_H
C 37	17	28.3	648	2	CL386389 RPT4_33
C 38	17	28.3	649	2	BB630176 BB630176
C 39	17	28.3	693	5	BU461127 603367904
C 40	17	28.3	737	7	CK308567 SB02046B2
C 41	17	28.3	810	8	BH301526 CH230-25C
C 42	17	28.3	931	9	CL076552 CH216-139
C 43	17	28.3	1082	6	CB988870 AGENCOURT
C 44	17	28.3	1107	2	BF185624 601814768
C 45	17	28.3	3723	3	AK079455 Mus muscu

ALIGNMENTS

RESULT 1
LOCUS CL055958/c 1012 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-81P12, Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL055958
VERSION CL055958.1 GI:40511871
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1012)
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E., and Wilson, R. A physical map of the xenopus tropicalis genome Unpublished (2003)
CONTACT: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submls@wustl.edu
Insert Length: 17500 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 56
High quality sequence stop: 378.
FEATURES
Location/Qualifiers
1..1012
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-81P12"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBA2.1; CHOR1-216 Xenopus tropicalis BAC library"

ORIGIN
Query Match 33.3%; Score 20; DB 9; Length 1012;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
35 TTGTGTATATACACAGT 54
|||||

No.	Score	Query Match	Length	DB	ID	Description
1	59	100	0	59	2	AAV23202
2	17	28.8		60	2	AAV23192
3	17	28.8		720	13	Lactococc
4	17	28.8		4835	13	AAc61637
5	17	28.8		24971	4	Bacterial
6	17	28.8		4AB11998	8	Ac61637 BONT/A ne
7	16	27.1		60	2	AAV23201
8	16	27.1		131	7	AAV23191
9	16	27.1		332	8	Lactococc
10	16	27.1		327	3	AAc61658
11	16	27.1		327	9	AAc61657
12	16	27.1		327	12	AAc61658
13	16	27.1		334	6	Novel hum
14	16	27.1		936	6	ABO98288
15	16	27.1		1344	8	ABO98288
16	16	27.1		1596	4	AAc61657
17	16	27.1		1736	13	AAc61658
18	16	27.1		2637	3	AAc61658
19	16	27.1		2637	3	AAc61658
20	16	27.1		3260	12	AAc61658

C	21	16	27.1	3316	12	ADO63072
C	22	16	27.1	3423	12	AD100600
C	23	16	27.1	3926	8	ABT17633
C	24	16	27.1	4154	8	ABT19487
C	25	16	27.1	5560	4	ABLO2800
C	26	16	27.1	14775	4	ABL17452
C	27	16	27.1	20035	10	ACC42662
C	28	16	27.1	50000	9	ADB16929
C	29	16	27.1	110000	2	AAV21209_14
C	30	16	27.1	110000	12	ADO79173_3
C	31	16	27.1	200620	12	ADT05677
C	32	16	27.1	349980	13	ADO56449
C	33	15	25.4	201	13	ADS89891
C	34	15	25.4	201	13	ADS89889
C	35	15	25.4	201	13	ADS39025
C	36	15	25.4	256	6	ABN24582
C	37	15	25.4	294	6	ABN20843
C	38	15	25.4	417	6	AB182559
C	39	15	25.4	423	12	ADN13201
C	40	15	25.4	430	4	AA117775
C	41	15	25.4	430	4	AA112646
C	42	15	25.4	436	4	AA110700
C	43	15	25.4	437	6	ABO99210
C	44	15	25.4	454	5	ABV14338
C	45	15	25.4	481	11	ABN80898

Adh132501 Human pro
Adh17775 Human bre
Aa11766 Human bre
Aa11070 Human bre
Abg9921 Human ORF
Abv14338 Human pro
Acrl80598 Breast Ca

ALIGNMENTS

RESULT 1

ID AAV23202 standard; DNA; 59 BP.

AC AAV23202

DT 28-JUL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cp37.

KW Lactococcus lactis; constitutional promoter; optimise; spacer; KW

238

OS Lactococcus lactis.
YY

FH	Key	Location/Qualifiers
ET	Promoter	4 59

```

E11      /*tag= 'u'
E12      /start= name= "ConstLitIna" promote=

```

XX
PN W09807846-A1

XX 26-FEB-1998
PD

XX 25-AUG-1997: 97WQ-DK000342

AA 23-AUG-1996: 96DK-00000886.

AA
PA (JENS/) . JENSEN P R

Jensen PR, Hammer

DR WPI; 1998-179062/16.

PT New artificial promc

PT expression of genes.

PS Claim 28; Page 55; 89pp; English

CC This is a Lactococcus

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 504.339 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657b-30

Perfect score: 59
Sequence: 1 CATCATTAAGTTATCTTCTC.....TATATATACCTTAGTACTGT 59

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl: 1: gb_ba: 2: gb_bhg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	6	BD006961 Artificial
2	33.9	270756	2	AC117317	AC117317 Rattus no
3	32.2	205235	2	AC151864	AC151864 Colobus g
4	32.2	237541	2	AC112564	AC112564 Rattus no
5	30.5	1072	3	AF438094	AF438094 Menarus t
6	30.5	36788	1	MSGB983CS	MSGB983CS
7	30.5	62643	8	SCB9871	SCB9871
8	30.5	100791	9	CMS01R18	CMS01R18
9	30.5	108799	9	AC090989	AC090989
10	30.5	108821	9	AL161791	AL161791 Homo sapi
11	30.5	110000	10	AB017261	AB017261 Human DNA
12	30.5	110000	10	AB014179	AB014179 Human DNA
13	30.5	123806	8	AP006417	AP006417
14	30.5	153987	5	EX005096	EX005096 Zebrafish
15	30.5	163987	5	EX005096	EX005096 Zebrafish
16	30.5	170380	10	AL808028	AL808028 Mouse DNA
17	30.5	174678	10	AC134449	AC134449 Mus muscu
18	30.5	175831	2	AC119435	AC119435 Mus muscu

20	18	30.5	176148	9	AC010798	AC010798 Homo sapi
21	18	30.5	178825	2	AP001203	AP001203 Homo sapi
22	18	30.5	181577	2	AC135119	AC135119 Mus muscu
23	18	30.5	190469	10	AC109212	AC109212 Mus muscu
24	18	30.5	211023	10	AC118734	AC118734 Mus muscu
25	18	30.5	215325	2	AC114171	AC114171 Rattus no
26	18	30.5	224622	2	AC097937	AC097937 Rattus no
27	18	30.5	228652	9	CNS01DMD	CNS01DMD
28	18	30.5	241073	2	AC095479	AC095479 Rattus no
29	18	30.5	243104	10	AL672244	AL672244 Mouse DNA
30	18	30.5	243300	2	AC098101	AC098101 Rattus no
31	18	30.5	245023	2	AC129467	AC129467 Rattus no
32	18	30.5	248926	2	AC111553	AC111553 Rattus no
33	18	30.5	256038	2	AC135932	AC135932 Rattus no
34	18	30.5	258648	2	AC130010	AC130010 Rattus no
35	18	30.5	264430	2	AC111510	AC111510 Rattus no
36	18	30.5	269203	1	MLEPRT10	MLEPRT10
37	18	30.5	270288	2	AC094065	AC094065 Rattus no
38	18	30.5	272405	2	AC094363	AC094363 Rattus no
39	18	30.5	293854	2	AC126646	AC126646 Rattus no
40	18	30.5	304725	2	AC122592	AC122592 Rattus no
41	18	30.5	305730	2	AC114347	AC114347 Rattus no
42	18	30.5	311823	2	AC095362	AC095362 Rattus no
43	17	28.8	60	6	BD006951	BD006951 Artificial
44	17	28.8	617	11	BV066004	BV066004 S212P6041
45	17	28.8	685	6	AR500569	AR500569 Sequence

ALIGNMENTS

RESULT 1
BD006961
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION
BD006961
VERSION
JP 2001503249-A/30.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE

1 (bases 1 to 59)
Hammerr,K. and Janssen,P.R.
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
Patent: JP 2001503249-A 30 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Lactococcus lactis
PN JP 2001503249-A/30
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
CC C12N15/09; C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter

FEATURES
source
1. 59
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 59; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 6.1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CATCATTAAGTTATCTTCAATTCGCCGGAATGTTGTATATATACCTTAGTACTGT 59

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1430.81 Seconds

(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-30

Sequence: 1 CATCATTAAGTTATCTTC.....TATTAATCCTTACTACTGT 59

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: gb_ests1.*
2: gb_ests2.*
3: gb_ests3.*
4: gb_ests4.*
5: gb_ests5.*
6: gb_ests6.*
7: gb_ests7.*
8: gb_ests8.*
9: gb_ests9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	32.2	612	6	CD332288 STRPUS37.
C 2	19	32.2	777	4	BG783071
C 3	19	32.2	1004	8	BZ574047 meh2.3483
C 4	18	30.5	413	6	CA649335 wrtln.pko
C 5	18	30.5	518	8	CC065628 fgma004f0
C 6	18	30.5	532	8	AO671093 HS_5432_B
C 7	18	30.5	615	7	CV498726 65019.1 M
C 8	18	30.5	649	9	CR189011 Reverse s
C 9	18	30.5	684	9	CR232325 Reverse s
C 10	18	30.5	686	9	CR232325 Reverse s
C 11	18	30.5	733	9	CL601256 OB_Ba007
C 12	18	30.5	752	7	CF660390
C 13	18	30.5	763	7	AG566765 Mus muscu
C 14	18	30.5	805	7	CF662568 CCLL03a08
C 15	18	30.5	881	9	CNS06JBX
C 16	18	30.5	958	9	CNS06JBX
C 17	18	30.5	1082	9	AG332609 Mus muscu
C 18	17	28.8	101	4	CG484260 OST18153
C 19	17	28.8	250	4	BI538048
C 20	17	28.8	252	2	AM329112
C 21	17	28.8	260	4	BI504201
C 22	17	28.8	288	4	BB505791
C 23	17	28.8	336	1	AJ780709
C 24	17	28.8	340	2	BF557306

25	17	28.8	359	7	CO727963	CO727963 UMC_bend
C 26	17	28.8	360	1	AU098102	AU098102 AU098102
C 27	17	28.8	363	5	BY517626	BY517626 BY517626
C 28	17	28.8	377	7	CO334089	CO334089 EK308851
C 29	17	28.8	383	5	BY406633	BY406633 BY406633
C 30	17	28.8	392	8	AO301517	AO301517 HS_2216_B
C 31	17	28.8	405	5	BY410233	BY410233 BY410233
C 32	17	28.8	408	5	BY455496	BY455496 BY455496
C 33	17	28.8	414	5	BY450700	BY450700 BY450700
C 34	17	28.8	445	4	BM499104	BM499104 947041F06
C 35	17	28.8	462	4	BG360766	BG360766 947041F06
C 36	17	28.8	464	6	CB536822	CB536822 771893_MA
C 37	17	28.8	472	6	CB536857	CB536857 771893_MA
C 38	17	28.8	474	6	CB225216	CB225216 10M30D04
C 39	17	28.8	502	7	CR453637	CR453637 CR453637
C 40	17	28.8	503	7	CR452333	CR452333 CR452333
C 41	17	28.8	510	7	CK542756	CK542756 rswb0.00
C 42	17	28.8	531	5	BO850788	BO850788 OGB13J02
C 43	17	28.8	543	8	AZ856980	AZ856980 2M0161B08
C 44	17	28.8	545	7	CO302498	CO302498 EK184304
C 45	17	28.8	555	4	BI444164	BI444164 da195f09

ALIGNMENTS

RESULT 1
LOCUS CD332288/c
DEFINITION STRPUS37.005111 Sea urchin embryo 20hr blastula stage cDNA library
MIMMPS37 Strongylocentrotus purpuratus cDNA clone
CD332288
CD332288.1 GI:34798810

ACCESSION CD332288
VERSION CD332288
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus

REFERENCE Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidae; Euechinoidae; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

1 (bases 1 to 612)
Poustka,A.J., Groth D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,
Reinhardt,R., Herwig,R., Panopoulou,G. and Lehnach,H.

Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters

Genome Res. 13 (12), 2736-2746 (2003)

Contact: Poustka AJ

laboratory 145, dept lehrach

Max-Planck-Institut fuer Molekulare Genetik

Innestr.63-73,-P-14195 Berlin, Germany

Tel: +49 30 8413 1235

Fax: +49 30 8413 1128

Email: poustka@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting
(ONF) to reduce sequencing redundancy. According to the ONF
procedure, clones that display the same hybridisation matrix with a
battery of 200 oligonucleotides are grouped into clusters. One
clone per ONF cluster is selected for sequencing. The size of each
cluster is an indicator of the frequency of a transcript in the
analysed library. The cluster size as well as the coordinates of
the other clones assigned to the same ONF cluster as the clone from
which the above EST is generated is available at the sea urchin
project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA
clones and filters are distributed via the Resource Center/Primary
Database of the German Human Genome Project (<http://www.rzpd.de>)

PCR primers

FORWARD: 5' CCCAGGCTTACCTTATCTCCGCTG 3' (M13SP) 5'-seq

BACKWARD: 5' GCTATTACGACGCTGCGAAGGCGATGTG 3' (M13SP) 3'-seq

Seq primer: 5'-CGGTCGGAATTCGCGGT-3' pSPORT3/86

High quality sequence stop: 612.

Location/Qualifiers

1..612
/organism="Strongylocentrotus purpuratus"

FEATURES
source

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-31

Perfect score: 60
Sequence: 1 CATRAGAGAACTTATCTCTG.....TATRATACATGACTACTGTT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2004as:*
12: geneeqn2004bs:*
13: geneeqn2004ds:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2 AAV23203	AAV23203 Lactococc
2	60	33.3	328	3 AAC00593	AAC00593 Human sec
3	20	33.3	330	2 AAX51602	AAX51602 Human sec
4	19	31.7	58	2 AAV23184	AAV23184 Lactococc
5	19	31.7	60	2 AAV23207	AAV23207 Lactococc
6	18	30.0	60	2 AAV23204	AAV23204 Lactococc
7	18	30.0	391	10 ADD26958	Add26958 Human adi
8	17	28.3	413	4 AAK65938	AAK65938 Human imm
9	17	28.3	976	4 AAF07896	AAF07896 Fusarium
10	17	28.3	3549	4 ABL10635	ABL10635 Drosophila
11	17	28.3	7308	4 ABL10634	ABL10634 Drosophila
12	17	28.3	11271	4 ABL06192	ABL06192 Drosophila
13	17	28.3	22951	11 ACN44558	ACN44558 Human gen
14	17	28.3	42373	10 ADC87696	ADC87696 Human mam
15	17	28.3	65047	11 ACN44020	ACN44020 Mouse gen
16	17	28.3	110000	6 ABA90521_03	ABA90521_03 of
17	17	28.3	110000	11 ACN44150_3	ACN44150_3 of
18	16	26.7	58	9 ABR98625	ABR98625 L. lactis
19	16	26.7	58	9 ACD13876	ACD13876 L. lactis
20	16	26.7	60	2 AAV23177	AAV23177 Lactococc

21	16	26.7	60	2 AAV23191	AAV23191 Lactococc
22	16	26.7	60	2 AAV23208	AAV23208 Lactococc
23	16	26.7	60	2 AAV23189	AAV23189 Lactococc
24	16	26.7	60	2 AAV23198	AAV23198 Lactococc
25	16	26.7	60	2 AAV23178	AAV23178 Lactococc
26	16	26.7	60	2 AAV23180	AAV23180 Lactococc
27	16	26.7	64	6 ABR98604	ABR98604 L. lactis
28	16	26.7	64	6 ACD13855	ACD13855 L. lactis
29	16	26.7	65	6 ABR98605	ABR98605 L. lactis
30	16	26.7	65	6 ACD13856	ACD13856 L. lactis
31	16	26.7	93	6 ABR98615	ABR98615 L. lactis
32	16	26.7	93	6 ACD13866	ACD13866 L. lactis
33	16	26.7	121	10 ADH92493	ADH92493 Human gen
34	16	26.7	121	10 ADH92494	ADH92494 Human gen
35	16	26.7	149	2 AAV02811	AAV02811 Human RHA
36	16	26.7	271	10 ABR82613	ABR82613 Bovine ear-
37	16	26.7	399	8 ABR54585	ABR54585 Bovine ES
38	16	26.7	466	6 ABR08596	ABR08596 Human leu
39	16	26.7	470	9 ACH26307	ACH26307 Human adu
40	16	26.7	510	2 AAT13640	AAT13640 ACNPV ORF
41	16	26.7	522	3 AAF09091	AAF09091 Fusarium
42	16	26.7	651	3 AAF13601	AAF13601 Aspergill
43	16	26.7	658	13 ADO48992	ADO48992 Novel can
44	16	26.7	710	6 ABO58231	ABO58231 Human col
45	16	26.7	755	6 ABR99221	ABR99221 Arabidops

ALIGNMENTS

RESULT 1
AAV23203
ID AAV23203 standard; DNA; 60 BP.
XX
XX AAV23203;
XX
DT 28-JUL-1998 (first entry)
XX
DE Lactococcus lactis constitutional promoter Cp38.
XX
XX Lactococcus lactis; constitutional promoter; optimise; spacer;
XX
XX artificial promoter library; gene expression; ds.
XX
XX Synthetic.
XX
XX Lactococcus lactis.
XX
OS Lactococcus lactis.
XX
XX Key Location/Qualifiers
XX
XX FT 4..60
XX
XX FT promoter /tag= a
XX
XX FT /standard_name= "Constitutional promoter"
XX
XX PN WO9807846-A1.
XX
XX PD 26-FEB-1998.
XX
XX PF 25-AUG-1997; 97WO-DK000342.
XX
XX PR 23-AUG-1996; 96DK-00000886.
XX
XX PA (JENS/) JENSEN P R.
XX
XX PI Jensen PR, Hammer K;
XX
XX DR WPI, 1998-179062/16.
XX
XX PT New artificial promoter libraries - containing consensus promoter
XX
XX PT sequences and variable spacers, used to generate promoters for optimising
XX
XX PT expression of genes.
XX
XX PS Claim 28; Page 55; 89pp; English.
XX
XX CC This is a Lactococcus lactis constitutional promoter sequence used in the
XX
XX CC construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-31

Sequence: 1 CATAGAGAGATTATTTCTTCTG.....TATATATCATGACTGTT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank1:*

1: gb_Da:*

2: gb_Ht:*

3: gb_In:*

4: gb_Om:*

5: gb_Ov:*

6: gb_Pat:*

7: gb_PN:*

8: gb_P1:*

9: gb_P2:*

10: gb_P3:*

11: gb_P4:*

12: gb_P5:*

13: gb_P6:*

14: gb_P7:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006962	BD006962 Artificialia
2	20	33.3	328	AX884728	AX884728 Sequence
3	20	33.3	328	BD024338	BD024338 Sequence
4	20	33.3	330	AX968753	AX968753 Sequence
5	20	33.3	330	BD073771	BD073771 5'EST of
6	20	33.3	176833	AC011994	AC011994 Homo sapi
7	20	33.3	209852	AC151363	AC151363 Gasterost
8	19	31.7	58	BD006943	BD006943 Artificialia
9	19	31.7	60	BD006966	BD006966 Artificialia
10	19	31.7	121171	AC105026	AC105026 Homo sapi
11	19	31.7	161162	AC022485	AC022485 Homo sapi
12	19	31.7	172348	AC121832	AC121832 Mus muscu
13	19	31.7	173071	AC009477	AC009477 Homo sapi
14	19	31.7	178095	CR753874	CR753874 Dario rer
15	19	31.7	194081	AC116526	AC116526 Mus muscu
16	19	31.7	202774	AC108851	AC108851 Mus muscu
17	19	31.7	209607	AC129581	AC129581 Mus muscu
18	19	31.7	227319	AC119894	AC119894 Mus muscu
19	19	31.7	283185	BX005447	BX005447 Zebrafish

20	18	30.0	60	BD006963	BD006963 Artificialia
21	18	30.0	391	AX806769	AX806769 Sequence
22	18	30.0	1943	AX124165	AX124165 Homo sapi
23	18	30.0	108484	AC004238	AC004238 Arabidops
24	18	30.0	141777	AC011738	AC011738 Homo sapi
25	18	30.0	147114	AP001939	AP001939 Homo sapi
26	18	30.0	158073	AC093756	AC093756 Homo sapi
27	18	30.0	159670	AC027626	AC027626 Homo sapi
28	18	30.0	162596	AC090417	AC090417 Homo sapi
29	18	30.0	164229	AC092671	AC092671 Homo sapi
30	18	30.0	167746	AC148606	AC148606 Gasterost
31	18	30.0	170022	AC120072	AC120072 Rattus no
32	18	30.0	170610	AC023935	AC023935 Homo sapi
33	18	30.0	174264	AL591373	AL591373 Mouse DNA
34	18	30.0	174711	AC026639	AC026639 Homo sapi
35	18	30.0	184892	AC016017	AC016017 Mus muscu
36	18	30.0	189534	AC123646	AC123646 Mus muscu
37	18	30.0	190708	CR381533	CR381533 Dario rer
38	18	30.0	196060	AC102143	AC102143 Mus muscu
39	18	30.0	197265	AL603702	AL603702 Mouse DNA
40	18	30.0	209591	AC113072	AC113072 Mus muscu
41	18	30.0	209643	AC079438	AC079438 Mus muscu
42	18	30.0	223724	AC114003	AC114003 Mus muscu
43	18	30.0	236402	AC109065	AC109065 Rattus no
44	18	30.0	237793	AC113832	AC113832 Rattus no
45	18	30.0	241726	AC131440	AC131440 Rattus no

ALIGNMENTS

RESULT 1
BD006962
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006962
ACCESSION
BD006962.1 GI:18635333
VERSION
JP 2001503249-A/31.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
1 (bases 1 to 60)
Hammer, K. and Janssen, P.R.

AUTHORS
Artificial promoter libraries for selected organisms and promoters
TITLE
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 31 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT
OS Lactococcus lactis
PN JP 2001503249-A/31
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PT KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandness: Double;
CC Topology: Linear;
FH Key
FT promoter

FEATURES
source
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CATAGAGAGATTATTTCTTACAGCTTGGCTTTGATATATATCATGACTGTT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds

(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-31

Perfect score: 60
Sequence: 1 CATAGAGAGTATTATCTTG.....TATPATCATGACTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	33.3	445	9	CL361142 RPCI44_34
2	20	33.3	570	9	CL385164 RPCI44_33
3	19	31.7	442	7	CN961352 8852_1002
4	19	31.7	628	9	CE335717 tigr-gss-
5	19	31.7	867	8	B2168391
6	18	30.0	307	9	CL316709 ZMMB0C046
7	18	30.0	415	7	HS9960 YR16E01.r1
8	18	30.0	429	2	AM195148 XM66E10.x
9	18	30.0	429	8	BH110154 RPCI-24-2
10	18	30.0	440	2	BE465718 hu34h08.x
11	18	30.0	453	2	BP176177 MYR6567a.
12	18	30.0	466	1	AT1218283 qb28g12.x
13	18	30.0	534	1	AA311111 EST181875
14	18	30.0	594	6	CD785518 EST656879
15	18	30.0	603	5	BX100105 BX100105
16	18	30.0	614	8	B2243274 CH230-272
17	18	30.0	754	9	CC829488 ZMMB0017
18	18	30.0	788	9	CL406797 ZMMB0041
19	18	30.0	898	9	CL065926 CH216-117
20	18	30.0	1054	2	BE748802 601572011
21	18	30.0	1166	2	CL024083 CH216-19A
22	18	30.0	1221	8	B2573764 msh2_3350
23	18	30.0	2031	3	AK044696 Mus muscu
24	17	28.3	156	8	AZ080173 RPCI-23-4

25	17	28.3	217	9	CR046864 Reverse s
26	17	28.3	257	1	AV070847
27	17	28.3	263	9	CE743019 tigr-gss-
28	17	28.3	292	9	CE038695 tigr-gss-
29	17	28.3	339	7	CO298913 EK174108.
30	17	28.3	356	9	CE153057 tigr-gss-
31	17	28.3	413	1	AA706239 ah28b1.s
32	17	28.3	447	7	R52514 YG81B02.s1
33	17	28.3	454	6	CD822486 BN25.045G
34	17	28.3	459	2	BE344464 EST113526
35	17	28.3	473	2	BE474293 sp61f03.y
36	17	28.3	474	9	BX997354 Reverse s
37	17	28.3	494	4	BG042217 su93b11.y
38	17	28.3	503	1	AA802413 GM05518.s
39	17	28.3	503	7	CF260824 NCST3409
40	17	28.3	523	1	AA255619 Z831C05.s
41	17	28.3	528	8	BH356815 CH230-6A1
42	17	28.3	543	7	CF797966 NCST3C62
43	17	28.3	548	5	BO081510 san25d01.
44	17	28.3	549	1	AI022239 ox01909.x
45	17	28.3	555	4	BI251438 602394667

ALIGNMENTS

RESULT 1
CL361142 445 bp DNA linear GSS 19-AUG-2004
LOCUS RPCI44_344U4.f RPCI-44 Sus scrofa genomic clone RPCI44_344U4,
DEFINITION genomic survey sequence.
ACCESSION CL361142
VERSION CL361142.1 GI:51413112
KEYWORDS GSS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

REFERENCE
AUTHORS Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Martron,B.M.,
Beever,J.B. and Schock,L.B.
TITLE Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
JOURNAL Through Comparative Genomics
COMMENT Unpublished (2004)
Other GSSs: RPCI44_344U4.x
Contact: Lawrence B. Schock
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schock@uiuc.edu

Clones are derived from the porcine BAC library RPCI-44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@chori.org).
Clones may be purchased from BACPAC Resources
(http://BACPACResources.chori.org). This work was undertaken as part
of the International Swine Genome Sequencing Consortium by
University of Illinois at Urbana Champaign, USA with funds provided
by grant No. AG2002-34480-11828 from USDA-CRSRS and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
Initiative)

Plate: 344 row: J column: 4
Seg primer: 17
Class: BAC ends.

FEATURES

source Location/Qualifiers
1..445
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="four pigs (bred: 37.5% Yorks Landrace and 25%
Meishan)"
/db_xref="taxon:9823"
/clone="RPCI44_344U4"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657b-32

Perfect score: 60

Sequence: 1 CATTCGGAAGTTATCTCTG.....TATTAATAGTACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001a:*\n5: geneseqn2001b:*\n6: geneseqn2002a:*\n7: geneseqn2002b:*\n8: geneseqn2003a:*\n9: geneseqn2003b:*\n10: geneseqn2003c:*\n11: geneseqn2003d:*\n12: geneseqn2004a:*\n13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23204
2	18	30.0	60	2	AAV23203
3	18	30.0	60	2	AAV23198
4	17	28.3	60	2	AAV23207
5	17	28.3	60	2	AAV23177
6	17	28.3	60	2	AAV23189
7	17	28.3	402	4	AA181365
8	17	28.3	5183	4	AAV04063
9	17	28.3	7364	12	ADG30623
10	17	28.3	29993	10	ADB37662
11	17	28.3	29993	10	ADB37660
12	17	28.3	38342	4	AA546745
13	17	28.3	38342	6	ABK31506
14	17	28.3	91823	10	ADL13497
15	16	26.7	58	2	AAV23184
16	16	26.7	58	6	ABK98625
17	16	26.7	58	9	ACD13876
18	16	26.7	60	2	AAV23191
19	16	26.7	60	2	AAV23208
20	16	26.7	60	2	AAV23178

21	16	26.7	60	2	AAV23180	AAV23180 Lactococc
22	16	26.7	64	6	ABK98604	ABK98604 L. lactis
23	16	26.7	64	9	ACD13855	ACD13855 L. lactis
24	16	26.7	65	6	ABK98605	ABK98605 L. lactis
25	16	26.7	65	6	ACD13856	ACD13856 L. lactis
26	16	26.7	93	6	ABK98615	ABK98615 L. lactis
27	16	26.7	93	6	ACD13866	ACD13866 L. lactis
28	16	26.7	261	3	ACD18503	ACD18503 Human sec
29	16	26.7	445	6	AB187329	AB187329 Human ova
30	16	26.7	522	3	AAH09091	AAH09091 Fusarium
31	16	26.7	600	12	ACH78367	ACH78367 Human gen
32	16	26.7	623	2	AAV41691	AAV41691 Nucleotid
33	16	26.7	651	3	AAFI3601	AAFI3601 Aspergill
34	16	26.7	738	12	ADL61118	ADL61118 Human pro
35	16	26.7	755	6	ABN99221	ABN99221 Arabidops
36	16	26.7	1015	2	AAK20572	AAK20572 Polynucle
37	16	26.7	1062	6	ABZ13652	ABZ13652 Arabidops
38	16	26.7	1062	6	ADG87654	ADG87654 A. thalia
39	16	26.7	1062	6	ADG87655	ADG87655 A. thalia
40	16	26.7	1062	8	ADA68053	ADA68053 Arabidops
41	16	26.7	1285	10	ADC87410	ADC87410 Human GPC
42	16	26.7	1305	3	AAK47953	AAK47953 Arabidops
43	16	26.7	1305	3	AAK51562	AAK51562 Arabidops
44	16	26.7	2118	3	AAK51464	AAK51464 Arabidops
45	16	26.7	8665	4	ABL03210	ABL03210 Drosophila

ALIGNMENTS

RESULT 1

ID: AAV23204 standard; DNA; 60 BP.

AAV23204;

28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cp39.

Lactococcus lactis; constitutional promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Synthetic.

Lactococcus lactis.

Key Location/Qualifiers

promoter 4..60 /tag= a

W09807846-A1. /standard_name= "Constitutional promoter"

26-FEB-1998. 97MO-DK000342.

25-AUG-1997; 96DK-00000886.

23-AUG-1996; 96DK-00000886.

(JENS/) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 55; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-32

Perfect score: 60
Sequence: 1 CATGCGAAGTTATCTTG.....TATAAAGTATAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

1: gb_ba:*
2: gb_hig:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	60	100.0	60	BD006963	BD006963 Artificial
2	60	33.3	93867	AL603747	AL603747 Zebrafish
3	20	33.3	208578	BX548050	BX548050 Zebrafish
4	19	31.7	2516	AK112417	AK112417 Clona Int
5	19	31.7	159314	AK121581	AK121581 Mus muscu
6	19	31.7	195899	AC123891	AC123891 Rattus no
7	19	31.7	222692	AC126962	AC126962 Rattus no
8	19	31.7	282700	AC130103	AC130103 Rattus no
9	18	30.0	60	BD006962	BD006962 Artificial
10	18	30.0	1877	PS081288	PS081288 Pleum saliv
11	18	30.0	17658	AC111945_5	AC111945_5 Continuation (6 of
12	18	30.0	43333	AC091064	AC091064 Homo sapi
13	18	30.0	100496	HS970D1	HS970D1 Human DNA
14	18	30.0	110060	AC098243_2	AC098243_2 Continuation (3 of
15	18	30.0	110060	AC111945_4	AC111945_4 Continuation (5 of
16	18	30.0	110000	AC111945_4	AC111945_4 Continuation (5 of
17	18	30.0	110000	AC121678_3	AC121678_3 Continuation (4 of
18	18	30.0	110000	AC010142	AC010142 Homo sapi
19	18	30.0	136462	9	AC010142

20	18	30.0	137333	5	AL935167	AL935167 Zebrafish
21	18	30.0	139409	2	AC133820	AC133820 Rattus no
22	18	30.0	145969	5	BX511153	BX511153 Zebrafish
23	18	30.0	160906	9	AL662874	AL662874 Human DNA
24	18	30.0	165242	9	AC009298	AC009298 Homo sapi
25	18	30.0	168704	2	AC151126	AC151126 Bos tauri
26	18	30.0	169947	2	AC109657	AC109657 Rattus no
27	18	30.0	173735	9	AC006518	AC006518 Homo sapi
28	18	30.0	180594	2	AC115770	AC115770 Mus muscu
29	18	30.0	182826	9	AC103976	AC103976 Homo sapi
30	18	30.0	185997	5	BX470139	BX470139 Zebrafish
31	18	30.0	191424	9	BS000127	BS000127 Pan trogl
32	18	30.0	193267	2	AC009409	AC009409 Homo sapi
33	18	30.0	194152	2	AC131741	AC131741 Mus muscu
34	18	30.0	196929	2	AC122634	AC122634 Rattus no
35	18	30.0	197638	10	AL669895	AL669895 Mouse DNA
36	18	30.0	197662	2	AC112689	AC112689 Mus muscu
37	18	30.0	215270	2	AC107800	AC107800 Mus muscu
38	18	30.0	215675	2	AC123264	AC123264 Rattus no
39	18	30.0	219510	2	AC125594	AC125594 Rattus no
40	18	30.0	233064	2	AC137218	AC137218 Rattus no
41	18	30.0	233878	2	AC098146	AC098146 Rattus no
42	18	30.0	233840	2	AC094367	AC094367 Rattus no
43	18	30.0	252150	2	AC098201	AC098201 Rattus no
44	18	30.0	264543	2	AC098212	AC098212 Rattus no
45	18	30.0	273729	2	AC106147	AC106147 Rattus no

ALIGNMENTS

RESULT 1	BD006963	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006963	Artificial promoter libraries for selected organisms and promoters derived from such libraries.			
DEFINITION	BD006963	Artificial promoter libraries for selected organisms and promoters derived from such libraries.			
ACCESSION	BD006963	GI:18635334			
VERSION	JP 2001503249-A/32				
KEYWORDS	JP 2001503249-A/32				
SOURCE	JP 2001503249-A/32				
ORGANISM	JP 2001503249-A/32				
REFERENCE	JP 2001503249-A/32				
AUTHORS	Hammer, K. and Janssen, P.R.				
TITLE	Artificial promoter libraries for selected organisms and promoters derived from such libraries				
JOURNAL	Patent: JP 2001503249-A 32 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
OS	Lactococcus lactis				
PN	JP 2001503249-A/32				
PD	13-MAR-2001				
PF	25-AUG-1997 JP 1998510287				
PR	23-AUG-1996 DK 0886/96				
PI	KALIN HAMMER, PETER RUDAR JANSSEN				
PC	CI2N15/09, CI2N15/00				
CC	Strandedness: Double;				
CC	Topology: Linear;				
CC	Key				
FT	promoter				
FEATURES	Location/Qualifiers				
source	1..60				
	/organism="Lactococcus lactis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN	100.0%; Score 60; DB 6; Length 60;				
Query Match	Best local similarity 100.0%; Pred. No. 2.6e-23;				
Matches	60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 CATGCGAAGTTATCTTGACAGTACGTTTACATGATATATAGTACTGTT 60				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-32
Perfect score: 60
Sequence: 1 CATGCGAAGTTATCTCTG.....TATATAGTATAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb est1:*
2: gb est2:*
3: gb_nuc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g8a1:*
9: gb_g8a2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	20	33.3	460	2	AM591004 X007D06.X
C 2	20	33.3	899	2	BF032856 60145696
C 3	20	33.3	1088	7	CK854279 15384 Stc
C 4	20	33.3	1126	8	A2130252 OSJNB010
C 5	19	31.7	404	8	AQ0807148 HS 3249 A
C 6	19	31.7	446	2	BE322759 NF047B051
C 7	19	31.7	533	8	AZ991259 2M0275A22
C 8	19	31.7	588	8	BE322833 100708280
C 9	19	31.7	608	8	BE626144 1007111F0
C 10	19	31.7	662	8	BE626144 1007111F0
C 11	19	31.7	726	9	AG170672 Pan trogl
C 12	19	31.7	802	9	CC633934 OGUIG04TV
C 13	19	31.7	967	9	CG461678 PRFXJ58TB
C 14	19	31.7	1003	9	CG461713 PRFXJ58TB
C 15	18	30.0	280	2	BE163245 BB163245
C 16	18	30.0	579	8	AZ778410 2M0013D12
C 17	18	30.0	626	2	BE523541 UR-R-C0-1
C 18	18	30.0	671	6	CA159481 SCUPR23C0
C 19	18	30.0	674	6	CA159393 SCUPR23C0
C 20	18	30.0	708	9	AG097174 Pan trogl
C 21	18	30.0	727	8	AQ479837 RPCI-11-2
C 22	18	30.0	774	8	B2119401 CH230-264
C 23	18	30.0	911	9	CR252276 Forward s
C 24	18	30.0	911	9	CR252276 Forward s

C 25	18	30.0	913	9	CR103634 Forward s
C 26	17	28.3	207	2	BE240051 EST404100
C 27	17	28.3	215	8	CC382565 PVRG006TD
C 28	17	28.3	217	9	CR046864 Reverse s
C 29	17	28.3	329	1	AL378049 McBR35F06
C 30	17	28.3	360	2	BE240039 EST404088
C 31	17	28.3	369	1	AL382437 McBC06G06
C 32	17	28.3	384	2	BE248652 NF022F09D
C 33	17	28.3	391	1	AL371466 McBA44D08
C 34	17	28.3	411	2	BE322892 NF025C10I
C 35	17	28.3	436	7	CR744006 CR744006
C 36	17	28.3	445	9	CL361142 RPCI44.34
C 37	17	28.3	454	1	AT794732 bd68d05.Y
C 38	17	28.3	454	2	BF520880 EST458353
C 39	17	28.3	456	1	AL381994 McBC04C03
C 40	17	28.3	457	2	BE706300 BB706300
C 41	17	28.3	459	2	BE821585 GM700014B
C 42	17	28.3	461	1	AL385553 McBC29C02
C 43	17	28.3	468	4	BG588193 EST490002
C 44	17	28.3	474	8	AQ986302 RPCI-23-3
C 45	17	28.3	475	6	BY602654 BY602654

ALIGNMENTS

RESULT 1
AM591004/c
LOCUS X007D06.X1 NCI CGAP U3 Homo sapiens cDNA clone IMAGE:2703251 3'
DEFINITION similar to SW:IDHA_NACPA_Q28480 ISOCITRATE DEHYDROGENASE [NAD]
SUBMITT ALPFA, MITOCHONDRIAL PRECURSOR ; mRNA sequence.

AM591004
AM591004.1 GI:7278150
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS NCI-CGAP
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:
image.llnl.gov/image/html/tresources.shtml
Seq primer: 40UP from Gibco
High quality sequence stop: 89.
Location/Qualifiers

FEATURES

source
1. 460
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2703251"
/tissue_type="poorly-differentiated endometrial
adenocarcinoma, 2 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP U3"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.45 Kb. Life Technologies catalog #: 11541-018"

ORIGIN
Query Match 33.3%; Score 20; DB 2; Length 460;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198,929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657b-33

Perfect score: 60
Sequence: 1 GATGTTTACTTATTCCTTG.....TATAATCGGATCCTTAAGA 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1980s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23205 Lactococc
2	19	31.7	59	2	AAV23190 Lactococc
3	18	30.0	60	2	AAV23213 Lactococc
4	17	28.3	60	2	AAV23197 Lactococc
5	17	28.3	60	2	AAV23208 Lactococc
6	17	28.3	60	2	AAV23214 Lactococc
7	16	26.7	58	2	AAV23178 Lactococc
8	16	26.7	58	2	AAV23184 Lactococc
9	16	26.7	59	2	AAV23209 Lactococc
10	16	26.7	60	2	AAV23211 Lactococc
11	16	26.7	60	2	AAV23188 Lactococc
12	16	26.7	60	2	AAV23182 Lactococc
13	16	26.7	60	2	AAV23212 Lactococc
14	16	26.7	60	2	AAV23180 Lactococc
15	16	26.7	207	6	ABK98627 L. lactis
16	16	26.7	207	6	ABK98627 L. lactis
17	16	26.7	242	6	ABK98617 L. lactis
18	16	26.7	242	6	ABK98617 L. lactis
19	16	26.7	651	3	AAV231601 Apeyrg11
20	16	26.7	727	6	ABQ44375 Oligonuc

21	16	26.7	727	6	ABQ44374	Abq44374 Oligonuc
22	16	26.7	957	6	ABZ32070	Abz32070 Candida a
23	16	26.7	1963	10	ADP93889	Adp93889 N. tilleyi
24	16	26.7	1963	10	ADP93889	Adp93889 N. tilleyi
25	16	26.7	2864	12	ADF50636	Adf50636 Nomuraea
26	16	26.7	3654	6	ABL40720	AbL40720 B. sphaer
27	16	26.7	6685	6	ABL32906	AbL32906 Human imm
28	16	26.7	9566	6	ABK98634	Abk98634 Vector pe
29	16	26.7	9566	9	ACD13885	AcD13885 L. lactis
30	16	26.7	11534	6	ABJ32343	AbJ32343 Human imm
31	16	26.7	13511	6	ABJ32281	AbJ32281 Human imm
32	16	26.7	13894	10	ADB59134	AdB59134 Toxicity-
33	16	26.7	14301	4	ABL02084	AbL02084 Drosophi
34	16	26.7	17967	6	ABJ33015	AbJ33015 Human imm
35	16	26.7	45158	11	ACN45178	AcN45178 Human gen
36	16	26.7	83888	11	ACN44208	AcN44208 Mouse gen
37	16	26.7	110000	6	ABA90521_00	AbA90521 Genomic s
38	16	26.7	185371	6	ABT10718	AbT10718 Human bre
39	16	26.7	272022	12	ADO97126	AdO97126 Human can
40	15	25.0	58	2	AAV23186	AAV23186 Lactococc
41	15	25.0	58	6	ABK98625	AbK98625 L. lactis
42	15	25.0	58	9	ACD13876	AcD13876 L. lactis
43	15	25.0	59	2	AAV23210	AAV23210 Lactococc
44	15	25.0	59	2	AAV23193	AAV23193 Lactococc
45	15	25.0	59	2	AAV23206	AAV23206 Lactococc

ALIGNMENTS

RESULT 1	AAV23205
ID	AAV23205 standard; DNA; 60 BP.
XX	AAV23205;
XX	28-JUL-1998 (first entry)
DT	Lactococcus lactis constitutional promoter Cp4.
XX	Lactococcus lactis; constitutional promoter; optimise; spacer;
XX	artificial promoter library; gene expression; ds.
KW	Synthetic.
XX	Lactococcus lactis.
OS	Lactococcus lactis.
XX	Key
XX	Location/Qualifiers
XX	4..60
XX	/*tag= a
XX	/standard_name= "Constitutional promoter"
XX	WO9807846-A1.
XX	26-FEB-1998.
XX	25-AUG-1997; 97WO-DX000342.
XX	23-AUG-1996; 96DX-00000886.
XX	(JENS/) JENSEN P R.
XX	Jensen PR, Hammer K;
XX	WPI; 1998-179062/16.
XX	New artificial promoter libraries - containing consensus promoter
XX	sequences and variable spacers, used to generate promoters for optimising
XX	expression of genes.
XX	Claim 28; Page 56; 89pp; English.
XX	This is a Lactococcus lactis constitutional promoter sequence used in the
XX	construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-33

Perfect score: 60
Sequence: 1 GATGTTTAGTTTATCTTG.....TATATCGGATCCTTAGA 60

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_by:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	6	BD006964 Artificia
2	20	33.3	170489	2	AC150513 Bos tauru
3	20	33.3	192540	2	AC137651 Bos tauru
4	20	33.3	209683	2	AC144596 Bos tauru
5	19	31.7	59	6	BD006949 Artificia
6	19	31.7	110000	1	AE017308 6
7	19	31.7	150400	10	AC122911 Mus muscu
8	19	31.7	160925	8	AP004332 Oryza sat
9	19	31.7	178046	2	AC137949 Mus muscu
10	19	31.7	179619	2	AC109259 Mus muscu
11	19	31.7	196351	2	AC138402 Mus muscu
12	18	30.0	60	6	BD006972 Artificia
13	18	30.0	1130	9	BC062784 Homo sapi
14	18	30.0	47302	9	AC012673 Homo sapi
15	18	30.0	78416	9	AL391669 Human DNA
16	18	30.0	112231	2	AC148651 Medicago
17	18	30.0	143229	2	AL161620 Homo sapi
18	18	30.0	145173	9	AC005599 Homo sapi
19	18	30.0	150295	2	AC120109 Homo sapi

20	18	30.0	150567	9	AC080078 Homo sapi
21	18	30.0	154879	2	AL450104 Human DNA
22	18	30.0	153269	9	AC120108 Homo sapi
23	18	30.0	164789	10	AC147185 Mus muscu
24	18	30.0	166155	2	AC093129 Papio anu
25	18	30.0	175572	2	AC025349 Homo sapi
26	18	30.0	180763	2	AC067826 Homo sapi
27	18	30.0	188700	2	AC021603 Homo sapi
28	18	30.0	217898	9	AC008498 Homo sapi
29	18	30.0	340350	2	AC112299 Rattus no
30	17	28.3	60	6	BD006937 Artificia
31	17	28.3	60	6	BD006956 Artificia
32	17	28.3	60	6	BD006967 Artificia
33*	17	28.3	39613	2	AC149316 Phakopsor
34	17	28.3	82518	2	AC092077 Oryza sat
35	17	28.3	110000	1	AE017308 2
36	17	28.3	110000	2	AC121380 1
37	17	28.3	110000	2	AC131176 2
38	17	28.3	126315	5	BX649476 Zebrafish
39	17	28.3	13612	9	AL353771 Human DNA
40	17	28.3	141405	10	AL354835 Mouse DNA
41	17	28.3	143683	9	AL356376 Human DNA
42	17	28.3	143839	9	AC004701 Homo sapi
43	17	28.3	144649	9	AC078591 Human DNA
44	17	28.3	145028	9	AC073875 Homo sapi
45	17	28.3	145028	9	AC073875 Homo sapi

ALIGNMENTS

RESULT 1
BD006964

LOCUS
BD006964

DEFINITION
Artificial promoter libraries for selected organisms and promoters

derived from such libraries.

ACCESSION
BD006964.1

GI:18635335

VERSION
JP 2001503249-A/33.

KEYWORDS
Lactococcus lactis

SOURCE
Lactococcus lactis

ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE
1 (bases 1 to 60)

Hammer, K. and Janssen, P.R.

AUTHORS
Artificial promoter libraries for selected organisms and promoters

derived from such libraries

JOURNAL
Patent: JP 2001503249-A 33 13-MAR-2001;

PETER RUDAR JANSSEN

COMMENT
OS Lactococcus lactis

PN JP 2001503249-A/33

PD 13-MAR-2001

PF 25-AUG-1997 JP 1998510287

PR 23-AUG-1996 DK 0886/96

PI KALIN HAMMER, PETER RUDAR JANSSEN

PC C12N15/09, C12N15/00

CC Strandedness: Double;

CC Topology: Linear;

FT Key Location/Qualifiers

FEATURES
location (4)..(60).

FT Promoter

source

1..60 location/Qualifiers
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;

Best Local Similarity 100.0%; Pred. No. 4e-26;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATGTTTAGTTTATCTTGACACCGATCGCGGTATATATCGGATCCTTAGA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds

(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-33

Perfect score: 1 GATGTTTACTTATCTCTG.....TATATCGGATCCTTAA 60

Sequence: 1 GATGTTTACTTATCTCTG.....TATATCGGATCCTTAA 60

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 segs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST.*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	19	31.7	612	CA387595	CA387595 669609 NC
2	19	31.7	646	AQ510836	AQ510836 nxb0048L
3	19	31.7	854	BX137003	BX137003 Danio rer
4	19	31.7	1383	AG394515	AG394515 Mus muscu
5	19	31.7	2650	AK036119	AK036119 Mus muscu
6	18	30.0	749	AG503191	AG503191 Mus muscu
7	18	30.0	913	BF528131	BF528131 Mus muscu
8	18	30.0	913	BF528131	BF528131 Mus muscu
9	18	30.0	913	BF528131	BF528131 Mus muscu
10	18	28.3	217	AA970556	AA970556 O09408.8
11	17	28.3	250	CG921685	CG921685 MBRPW77F
12	17	28.3	304	BB193512	BB193512 BB193512
13	17	28.3	325	BB393516	BB393516 BB393516
14	17	28.3	326	AI430702	AI430702 mcs59h10.Y
15	17	28.3	446	CG126683	CG126683 PIMY04TB
16	17	28.3	466	AQ495913	AQ495913 HS_5056_A
17	17	28.3	466	AQ724742	AQ724742 HS_5402_A
18	17	28.3	517	AZ065720	AZ065720 RCT-23-4
19	17	28.3	518	BX616883	BX616883 BX616883
20	17	28.3	529	AQ595322	AQ595322 HS_5422_B
21	17	28.3	529	AQ595322	AQ595322 HS_5422_B
22	17	28.3	587	CK450842	CK450842 903947_MA
23	17	28.3	591	BZ717691	BZ717691 PUCV34TD
24	17	28.3	600	CR532508	CR532508 CR532508

C 25	17	28.3	613	9	DR23B35	AL966013 Danio rer
C 26	17	28.3	621	2	BB658648	BB658648 BB658648
C 27	17	28.3	628	9	CE335717	CE335717 Cigr-g88-
C 28	17	28.3	629	2	BB661991	BB661991 BB661991
C 29	17	28.3	635	5	CE120196	CE120196 Cigr-g88-
C 30	17	28.3	652	5	BX616883	BX616883 BX616883
C 31	17	28.3	654	8	BH165823	BH165823 BHTSS91TF
C 32	17	28.3	675	8	BH978524	BH978524 Odf89f08.
C 33	17	28.3	683	7	CN625364	CN625364 cae01a11.
C 34	17	28.3	708	8	BH988921	BH988921 oeh39f08.
C 35	17	28.3	728	8	CR330935	CR330935 Medicago
C 36	17	28.3	737	2	BB184575	BB184575 BB184575
C 37	17	28.3	781	5	BU144062	BU144062 Tor1634 G
C 38	17	28.3	819	7	CN201575	CN201575 Tor1634 G
C 39	17	28.3	836	9	CR304111	CR304111 Medicago
C 40	17	28.3	840	8	BH158246	BH158246 ENTOS46TR
C 41	17	28.3	847	8	AZ669235	AZ669235 ENTMS77RF
C 42	17	28.3	917	9	CL902633	CL902633 CSHC2008
C 43	17	28.3	920	8	AZ528690	AZ528690 ENTBU51TR
C 44	17	28.3	967	7	W35992	W35992 mcs59h10.X1
C 45	17	28.3	989	9	CG105337	CG105337 PUFZA70TD

ALIGNMENTS

RESULT 1
CA387595
LOCUS
DEFINITION 669609 NCCOWA 1RT Oncorhynchus mykiss cDNA clone 1RT13B12_D_A06
VERSION CA387595
KEYWORDS 5', mRNA sequence.
SOURCE CA387595.1 GI:24716016
ORGANISM EST.
Oncorhynchus mykiss (rainbow trout)
Oncorhynchus mykiss
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 612)
Rexroad, C.E. 3rd, Lee, Y., Kaele, J.W., Karamycheva, S., Brown, G.,
Koop, B., Gahr, S.A., Palti, Y. and Quackenbush, J.
Sequence analysis of a rainbow trout cDNA library and creation of a
gene index
JOURNAL CytoGenet. Genome Res. 102 (1-4), 347-354 (2003)
COMMENT Contact: Rexroad CE
USDA, ARS, National Center for Cool and Cold Water Aquaculture
11876 Leetown Road, Kearneysville, WV 25430, USA
Tel: 304 724 8340 X2129
Fax: 304 725 0351
Email: crexroad@nccowa.ars.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim alt option. Vector identified by
cross_match v0.990329.
Seq primer: AGCGATACATTTTCACACAGA.
Location/Qualifiers
1..612
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="1RT13B12_D_A06"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="NCCOWA 1RT"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from brain, gill, liver,
spleen, muscle, and kidney."

ORIGIN

Query Match 31.7%; Score 19; DB 6; Length 612;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 / Search time 195.613 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657b-34

Perfect score: 59

Sequence: 1 CATAGAACAGTTATCTTG.....ATATATAGCCAGTACTGTT 59

Scoring table: Oligo_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	2	AAV23206
2	20	33.9	60	2	AAV23194
3	19	32.2	58	2	AAV23186
4	19	32.2	23973	12	AD182449
5	18	30.5	1189	4	AAH47068
6	18	30.5	3693	12	ADN37297
7	18	30.5	11777	10	AD123597
8	18	30.5	12079	6	ABV74271
9	18	30.5	12079	6	ABO76793
10	18	30.5	12079	13	ADR49334
11	18	30.5	12085	6	ABV74270
12	18	30.5	12085	6	ABO76792
13	18	30.5	12085	6	ABO76793
14	18	30.5	12093	6	ABV74269
15	18	30.5	12093	6	ABO76791
16	18	30.5	12093	13	ADR49332
17	18	30.5	12241	6	AAO36732
18	18	30.5	12241	6	ABO73049
19	18	30.5	12241	11	AD005418
20	18	30.5	12982	4	AAE6432

21	18	30.5	13002	6	ABV74272
22	18	30.5	13002	6	ABO76794
23	18	30.5	13002	13	ADR49335
24	18	30.5	13718	12	ADH56070
25	18	30.5	13737	3	AAH54212
26	18	30.5	13905	6	ABV74273
27	18	30.5	13905	6	ABO76795
28	18	30.5	13905	13	ADR49336
29	18	30.5	14113	3	AAH39651
30	18	30.5	14113	3	AAH39651
31	18	30.5	14113	4	AAH25853
32	18	30.5	14446	6	AAH17548
33	18	30.5	14603	6	AAH36966
34	18	30.5	15294	13	ADH33243
35	18	30.5	15430	6	ABV74274
36	18	30.5	15430	6	ABO76796
37	18	30.5	15643	6	AAH36963
38	18	30.5	15739	13	ADR03845
39	18	30.5	15739	13	ADR03925
40	18	30.5	16103	13	ADR03904
41	18	30.5	16103	13	ADR03984
42	18	30.5	16179	6	AAH36962
43	18	30.5	16245	13	ADR03878
44	18	30.5	16245	13	ADR03958
45	18	30.5	16954	13	ADR03886

ALIGNMENTS

RESULT 1	AAV23206	standard; DNA, 59 BP.
ID	AAV23206	
XX	AAV23206;	
AC	28-JUL-1998	(first entry)
XX		
DT		
XX		
DE	Lactococcus lactis constitutional promoter Cp40.	
XX		
XX	Lactococcus lactis; constitutional promoter; optimise; spacer;	
KW	artificial promoter library; gene expression; ds.	
XX		
OS	Synthetic.	
XX	Lactococcus lactis.	
XX		
XX	Key	Location/Qualifiers
FT	Promoter	4..60
FT		/*tag= a
XX		/standard_name= "Constitutional promoter"
XX		
PN	W09607846-AI.	
XX		
PD	26-FEB-1998.	
XX		
XX	25-AUG-1997;	97WO-DX000342.
XX		
XX	23-AUG-1996;	96DK-00000886.
XX		
XX	(JENS) JENSEN P R.	
XX		
XX	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
XX	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
XX	Claim 28; Page 56; 89pp; English.	
PS		
XX	This is a Lactococcus lactis constitutional promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 504.339 Seconds

(Without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-34

Perfect score: 59

Sequence: 1 CATAGAACGTTATCTGCTTG.....ATTTATACCCAGTACTGTT 59

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ALIGNMENTS

Query Match	Score	DB 6:	Length	59:	Best Local Similarity	Pred. No.	2.3e-22:	Indels	Gaps	0:
Matches	59:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:	

Database :

GenBml:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sbs:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	59	100.0	59	6	BD006965	BD006965 Artificial
2	20	33.9	60	6	BD006953	BD006953 Artificial
3	20	33.9	36404	3	AF106592	AF106592 Caenorhab
4	20	33.9	110000	8	CR382137_14	Continuation (15 o
5	20	33.9	188406	10	AL713914	AL713914 Mouse DNA
6	20	33.9	200704	9	AC099781	AC099781 Homo sapi
7	19	32.2	58	6	BD006945	BD006945 Artificial
8	19	32.2	39578	2	AC101097	AC101097 Mus muscu
9	19	32.2	73632	10	AC098685	AC098685 Mus muscu
10	19	32.2	110000	2	AC116234_2	Continuation (3 of
11	19	32.2	146782	2	AC127599	AC127599 Rattus no
12	19	32.2	163364	2	AC116241	AC116241 Rattus no
13	19	32.2	187532	9	AC006333	AC006333 Homo sapi
14	19	32.2	213870	10	AL844180	AL844180 Mouse DNA
15	19	32.2	224818	2	AC111839	AC111839 Rattus no
16	19	32.2	250178	2	AC118121	AC118121 Rattus no
17	18	30.5	1189	6	AX212282	AX212282 Sequence
18	18	30.5	3696	1	AF065243	AF065243 Agrobacte
19	18	30.5	3807	1	AE009438	AE009438 Agrobacte

RESULT 1
BD006965
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS
PN
PD
PF
PI
PC
CC
FH
FT
FEATURES
ORIGIN
Query Match
Best Local Similarity
Matches

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 / Search time 1430.81 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-34

Perfect score: 59
Sequence: 1 CATAGAACAGTATTCTTCTG.....ATATATAGCCACTACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gest1: *
9: gb_gest2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	20	33.9	437	7 N58914	N58914 yy61e12.g1
2	20	33.9	804	9 CNS04C60	AL284049 Tetracodon
3	20	33.9	1111	8 CC277811	CC277811 CH261-30K
4	18	30.5	288	1 AV094766	AV094766 AV094766
5	18	30.5	413	6 BV654942	BV654942 BV654942
6	18	30.5	550	9 CE199559	CE199559 C1gr-G98-
7	18	30.5	554	6 BH749677	BH749677 SALX 0286
8	18	30.5	561	8 BH747392	BH747392 SALX 0169
9	18	30.5	569	8 BH747092	BH747092 SALX 0100
10	18	30.5	578	8 BZ231363	BZ231363 CH230-380
11	18	30.5	583	8 BZ280238	BZ280238 CH230-297
12	18	30.5	613	9 CE524504	CE524504 C1gr-G98-
13	18	30.5	626	9 CC774010	CC774010 CH240 40F
14	18	30.5	654	7 CR772226	CR772226 DKF2P468E
15	18	30.5	717	7 CF472227	CF472227 RTD51.5 A
16	18	30.5	801	8 AO741731	AO741731 HS_556C_B
17	18	30.5	864	8 AO739419	AO739419 HS_5387_B
18	17	28.8	179	1 AL451605	AL451605 mus1e12.x
19	17	28.8	186	1 AA184870	AA184870 mus1e12.x
20	17	28.8	187	1 AV256985	AV256985 AV256985
21	17	28.8	192	1 AL627071	AL627071 mus1e12.y
22	17	28.8	206	5 BX527779	BX527779 BX527779
23	17	28.8	208	1 AU038817	AU038817 AU038817
24	17	28.8	233	4 BM218766	BM218766 C0915B09-

C 25	17	28.8	236	1 AV233619	AV233619
C 26	17	28.8	263	2 BB043394	BB043394
C 27	17	28.8	270	1 AV297539	AV297539
C 28	17	28.8	274	1 AV028171	AV028171
C 29	17	28.8	284	2 BB354784	BB354784
C 30	17	28.8	286	2 BB413979	BB413979
C 31	17	28.8	286	2 BB944623	BB944623
C 32	17	28.8	290	2 BB397484	BB397484
C 33	17	28.8	293	2 BB440039	BB440039
C 34	17	28.8	300	1 AV030326	AV030326
C 35	17	28.8	330	9 CR307025	CR307025
C 36	17	28.8	341	6 CB93956	CB93956
C 37	17	28.8	343	8 A2721448	A2721448
C 38	17	28.8	360	1 AJ485148	AJ485148
C 39	17	28.8	360	1 AJ485149	AJ485149
C 40	17	28.8	360	1 AJ485151	AJ485151
C 41	17	28.8	365	6 BY681332	BY681332
C 42	17	28.8	368	1 AU257848	AU257848
C 43	17	28.8	368	5 BY415434	BY415434
C 44	17	28.8	369	5 BY395944	BY395944
C 45	17	28.8	372	2 BB794085	BB794085

ALIGNMENTS

RESULT 1
N58914
LOCUS
DEFINITION
yy61e12.g1 Soares multiple sclerosis 2NBHSP Homo sapiens cDNA
clone IMAGE:278062 3', mRNA sequence.
N58914
N58914.1 GI:1202804
EST.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 437)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNC, contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mod.RBGA+T
High quality sequence scop: 1.
Location/Qualifiers
1. 437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3896438"
/db_xref="taxon:9606"
/clone="IMAGE:278062"
/sex="male"
/tissue-type="multiple sclerosis lesions"
/dev stage="Age 46"
/lab_host="PHIOB (ampicillin resistant)"
/note="lib=Soares multiple sclerosis 2NBHSP"
/note="Vector: pTR73D (Pharmacia) with a modified
polylinker V type: phagemid; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TGTTCACATCGAAGTGGAGCGCGCATTTTTTTTTTTTTTT 3']

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
              (without alignments)
              1785.485 Million cell updates/sec
```

Title: US-09-242-657B-35

Sequence: 1 CATCCGCAAGTTATTCTTG.....TATATAAGTTAAGTACTGTT 60

Scoring table: OLIGO_NUC

Searched: 4390206 seqs, 2959870667 residues

Word size :

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length:	0
------------------------	---

Maximum DB seq length: 200000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04 : *

```

1:  gensseqn190bs: *
2:  gensseqn190bs: *
3:  gensseqn2001as: *
4:  gensseqn2001as: *
5:  gensseqn2002bs: *
6:  gensseqn2002bs: *
7:  gensseqn2002bs: *
8:  gensseqn2003as: *
9:  gensseqn2003bs: *
10: gensseqn2003cs: *
11: gensseqn2003ds: *
12: gensseqn2004as: *
13: gensseqn2004bs: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60	100.0	60	2	AAV23207	AAV23207 Lactococc
2	20	33.3	651	3	AAFI3601	AAFI3601 Aspergill
3	19	31.7	60	2	AAV23203	AAV23203 Lactococc
4	18	30.0	58	2	AAV23184	AAV23184 Lactococc
5	17	28.3	60	2	AAV23204	AAV23204 Lactococc
6	17	28.3	60	2	AAV23180	AAV23180 Lactococc
7	17	28.3	809	4	AAI94629	AAI94629 Human neu
8	17	28.3	109586	11	ACMA3994	ACMA3994 Human ge
9	17	28.3	123331	12	ADQ87852	ADQ87852 Human ca
10	16	26.7	58	6	ABK98625	ABK98625 L. lactis
11	16	26.7	58	6	ACDI1876	ACDI1876 L. lactis
12	16	26.7	60	2	AAV23177	AAV23177 Lactococc
13	16	26.7	60	2	AAV23191	AAV23191 Lactococc
14	16	26.7	60	2	AAV23208	AAV23208 Lactococc
15	16	26.7	60	2	AAV23189	AAV23189 Lactococc
16	16	26.7	60	2	AAV23198	AAV23198 Lactococc
17	16	26.7	60	2	AAV23178	AAV23178 Lactococc
18	16	26.7	64	6	ABK98604	ABK98604 L. lactis
19	16	26.7	64	9	ACDI1855	ACDI1855 L. lactis
20	16	26.7	65	6	ABK98605	ABK98605 L. lactis

C	21	16	26.7	65	9	ACD13856	Adh1856 L. lactis
C	22	16	26.7	93	6	ABK96615	Adh96615 L. lactis
C	23	16	26.7	93	9	ACD13866	Adh1866 L. lactis
C	24	16	26.7	262	4	AAK65318	AAK65318 Human imm
C	25	16	26.7	262	4	AAK65319	AAK65319 Human imm
C	26	16	26.7	399	8	ABX54585	Bovine ES
C	27	16	26.7	508	4	AAI18176	Adh18176 Probe #81
C	28	16	26.7	508	4	ABA63146	Human foe
C	29	16	26.7	508	4	AAI43193	Adh43193 Probe #11
C	30	16	26.7	508	4	ABA30396	Adh30396 Probe #88
C	31	16	26.7	508	4	AAK37348	AAK37348 Human bon
C	32	16	26.7	508	4	AAK11581	AAK11581 Human bra
C	33	16	26.7	508	4	ABK37021	Human liv
C	34	16	26.7	755	6	ABK99221	Adh99221 Arabidops
C	35	16	26.7	1662	6	ABE13652	ABz13652 Arabidops
C	36	16	26.7	1662	6	ADG87654	ADG87654 A. thalia
C	37	16	26.7	1062	6	ADG87655	ADG87655 A. thalia
C	38	16	26.7	1062	8	ADG68053	ADG68053 Arabidops
C	39	16	26.7	1305	3	AAK47953	AAK47953 Arabidops
C	40	16	26.7	1305	3	AAK51562	AAK51562 Arabidops
C	41	16	26.7	1350	6	ADN22119	Adh22119 Human TST
C	42	16	26.7	1362	2	AAV99906	AAV99906 Fragment
C	43	16	26.7	2118	3	AAK51464	AAK51464 Arabidops
C	44	16	26.7	2454	12	ADP28007	Adp28007 Human sec
C	45	16	26.7	9417	6	ABK98632	Abk98632 Vector pb

ALIGNMENTS

```

RESULT 1
AAV23207
ID AAV23207 standard; DNA; 60 BP.
....

```

DT 28-JUL-1998 (first entry)

DE Lactococcus lactis constitutional promoter Cp41.

KW Lactococcus lactis; constitutional promoter; optimise; spacer;
KW artificial promoter library; gene expression; ds.

05 Synthetic.
05 Lactococcus lactis.

Key promoter	Location/Qualifiers
FT	4. .60

FT	vv	standard_name=	"Constitutional promoter"
		/standard_name=	

PN WO9807846-A1

PD 26-FEB-1998

PF 25-AUG-1997; 97WO-DK000342.

PR 23-AUG-1996; 96DK-00000886.

PA (JENSEN) JENSEN P 1

PI Jensen PR, Hammer K,

DR WPI; 1998-179062/16.
XX

PT New artificial promc
DT sequences and variab

PT expression of genes.
XX

PS Claim 28; Page 57; 89pp; English
XX

CC This is a *Lactococcus lactis* constitutional promoter sequence used in the
CC construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-35

Perfect score: 60
Sequence: 1 CATCCGCACTTATCTCTG.....TATATAGTAACTACTGTT 60

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Genembl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006966	BD006966 Artificia
2	19	31.7	60	BD006962	BD006962 Artificia
3	19	31.7	184537	AC126079	AC126079 Rattus no
4	18	30.0	58	BD006943	BD006943 Artificia
5	18	30.0	1829	BC075654	BC075654 Mus muscu
6	18	30.0	123556	AC149741	AC149741 Bos tauru
7	18	30.0	152177	AL935114	AL935114 Zebrafish
8	18	30.0	184892	AC016017	AC016017 Mus muscu
9	18	30.0	196124	AC141387	AC141387 Rattus no
10	18	30.0	203595	CR385054	CR385054 Danio rer
11	18	30.0	206733	AC110553	AC110553 Mus muscu
12	18	30.0	207788	AC095677	AC095677 Rattus no
13	18	30.0	209643	AC079438	AC079438 Mus muscu
14	18	30.0	215147	AC119806	AC119806 Mus muscu
15	18	30.0	220464	AC115409	AC115409 Rattus no
16	18	30.0	245753	AC097600	AC097600 Rattus no
17	18	30.0	247786	AC098752	AC098752 Rattus no
18	18	30.0	288968	AC115432	AC115432 Rattus no
19	17	28.3	60	BD006939	BD006939 Artificia

20	17	28.3	60	BD006963	BD006963 Artificia
21	17	28.3	433	AY120570	AY120570 Unculture
22	17	28.3	433	AY120572	AY120572 Unculture
23	17	28.3	433	AY120600	AY120600 Unculture
24	17	28.3	438	AY120569	AY120569 Unculture
25	17	28.3	809	BD018466	BD018466 Novel gen
26	17	28.3	809	BD098404	BD098404 Novel gen
27	17	28.3	2499	AF182946	AF182946 Rattus no
28	17	28.3	2568	AF057157	AF057157 Mus muscu
29	17	28.3	9281	AF547990	AF547990 Mus muscu
30	17	28.3	59362	AC100236	AC100236 Mus muscu
31	17	28.3	99832	AL837507	AL837507 Mouse DNA
32	17	28.3	106839	HS60631	Z94722 Human DNA s
33	17	28.3	110000	AC096323_0	AC096323 Rattus no
34	17	28.3	112182	AC141177	AC141177 Rattus no
35	17	28.3	124090	AC096657	AC096657 Homo sapi
36	17	28.3	127234	AL139155	AL139155 Homo sapi
37	17	28.3	140156	AC027261	AC027261 Homo sapi
38	17	28.3	143710	AP002899	AP002899 Oryza sat
39	17	28.3	146175	AL663084	AL663084 Mouse DNA
40	17	28.3	155263	AC053517	AC053517 Homo sapi
41	17	28.3	160141	AP003142	AP003142 Oryza sat
42	17	28.3	161193	AC011033	AC011033 Homo sapi
43	17	28.3	161298	AP002858	AP002858 Homo sapi
44	17	28.3	162333	AC144673	AC144673 Mus muscu
45	17	28.3	164741	BX005455	BX005455 Zebrafish

ALIGNMENTS

RESULT 1
BD006966
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006966
ACCESSION
BD006966.1 GI:16635337
VERSION
UP 2001503249-A/35.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
1 (bases 1 to 60)

AUTHORS
Hammer, K. and Janssen, P.R.
TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
Patent: JP 2001503249-A 35 13-MAR-2001;

JOURNAL
PETER RUDAR JANSSEN

COMMENT
OS Lactococcus lactis

PN UP 2001503249-A/35
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96

PI KALIN HAMMER, PETER RUDAR JANSSEN
CC C12N15/09, C12N15/00
PC Strandedness: Double;

CC Topology: Linear;

CC Key
FH Key
FT promoter

Location/Qualifiers
(4). (60).

FEATURES

source
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.4e-24;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATCCGCACTTATCTCTGACGAGTGAATGAGCGTGTATTAATAGTAACTACTGTT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-35

Perfect score: 60
Sequence: 1 CATCCGCAAGTTATCTGTG.....TATATAGTAAGTAAGTCTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	33.3	435	7	CO756473
2	20	33.3	644	7	CO151372
3	20	33.3	673	7	CO152362
4	18	30.0	418	9	CR046864
5	18	30.0	498	9	CE535916
6	18	30.0	565	2	BF427975
7	18	30.0	656	2	CO044252
8	18	30.0	729	4	BI657225
9	18	30.0	741	2	AW906037
10	17	28.3	285	9	CG632062
11	17	28.3	340	6	CB694443
12	17	28.3	445	4	BM389778
13	17	28.3	456	2	CR476358
14	17	28.3	484	2	BF587441
15	17	28.3	499	5	BQ560041
16	17	28.3	520	9	CG636000
17	17	28.3	520	9	CG642667
18	17	28.3	549	8	CG655759
19	17	28.3	551	8	AZ664444
20	17	28.3	573	7	CK682245
21	17	28.3	576	8	AO520938
22	17	28.3	612	6	CA387595
23	17	28.3	627	7	CN383241
24	17	28.3	636	9	CE563332

c 25	17	28.3	647	6	CA092216	CA092216
c 26	17	28.3	656	4	BM25023	BM25023
c 27	17	28.3	661	9	CG842276	CG842276
c 28	17	28.3	663	9	LBAP042E03	LBAP042E03
c 29	17	28.3	696	9	AG101542	AG101542
c 30	17	28.3	706	8	AO853572	AO853572
c 31	17	28.3	708	6	CB202605	CB202605
c 32	17	28.3	723	9	CG848135	CG848135
c 33	17	28.3	727	9	BX216300	BX216300
c 34	17	28.3	764	7	CO009449	CO009449
c 35	17	28.3	767	7	CO003923	CO003923
c 36	17	28.3	770	9	CO013254	CO013254
c 37	17	28.3	779	9	CG891143	CG891143
c 38	17	28.3	780	6	CA513507	CA513507
c 39	17	28.3	800	9	CG67818	CG67818
c 40	17	28.3	807	9	CR323339	CR323339
c 41	17	28.3	825	8	BZ799167	BZ799167
c 42	17	28.3	825	5	BU423415	BU423415
c 43	17	28.3	844	9	CG953853	CG953853
c 44	17	28.3	935	7	CF581300	CF581300
c 45	17	28.3	958	8	BZ502054	BZ502054

ALIGNMENTS

RESULT 1
LOCUS CO756473
DEFINITION Mdfrc3047d01.y1 Mdfrc3047d01.y1 Mdfrc3047d01.y1 Mdfrc3047d01.y1
similar to SW-TLPH_ARATH P50699 THAUMATIN-LIKE PROTEIN PRECURSOR.
; RNA sequence.
CO756473
CO756473.1 GI:50891720

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Malus x domestica (cultivated apple)
EST.
CO756473.1 GI:50891720

REFERENCE
AUTHORS
Korban, S., Vodka, L., Liu, L., Gasic, R., Gonzales, O., Hernandez, A.,
Alwinckle, H., Malnoy, M., Carroll, N., Goldsbrough, P., Ory, K.,
Clifton, S., Page, D., Marre, M., Hillier, L., Martin, J., Wylie, T.,
Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I.,
Tsagareishvili, R., Kennedy, S., Waterson, R., and Watson, R.
Apple Functional Genomics grant - NSF 0321702
Unpublished (2004)
Contact: Schuyler S. Korban

TITLE
JOURNAL
COMMENT
Apple Functional Genomics grant - NSF 0321702
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

FEATURES
Source
High quality sequence atp: 419.
Location/Qualifiers
1..435
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="Mdfrc3047d01"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdfrc"
/note="Vector: pBluescript II SK (+); Site: 1: NotI;
Site 2: EcoRII; Total RNA was extracted separately from
each stage [young fruitlet (<1cm), young fruitlet (1 cm
dia.), young fruitlet (12cm dia.), maturing fruit I,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-36

Perfect score: 1 CATTCTGTAAGTTTACTTCTG.....TATTAATAAATACTACTGTT 60

Sequence: 1 CATTCTGTAAGTTTACTTCTG.....TATTAATAAATACTACTGTT 60

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:**

1: geneseq1980s:**
2: geneseq1990s:**
3: geneseq1990s:**
4: geneseq2000s:**
5: geneseq2001as:**
6: geneseq2002as:**
7: geneseq2002bs:**
8: geneseq2003as:**
9: geneseq2003bs:**
10: geneseq2003cs:**
11: geneseq2003ds:**
12: geneseq2004as:**
13: geneseq2004bs:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	100.0	60	AAV23208	AAV23208 Lactococc
2	19	31.7	3059	AD035404	AD035404 Novel mou
3	18	30.0	1334	ADR61049	ADR61049 Cotton CD
4	17	28.3	59	AAV23190	AAV23190 Lactococc
5	17	28.3	60	AAV23180	AAV23180 Lactococc
6	17	28.3	60	AAV23205	AAV23205 Lactococc
7	17	28.3	349	AAV23150	AAV23150 Novel hum
8	17	28.3	651	AAV23150	AAV23150 Novel hum
9	17	28.3	4044	AB102942	AB102942 Drosophila
10	17	28.3	57137	AAV23211	AAV23211 Lactococc
11	16	26.7	60	AAV23203	AAV23203 Lactococc
12	16	26.7	60	AAV23207	AAV23207 Lactococc
13	16	26.7	60	AAV23182	AAV23182 Lactococc
14	16	26.7	60	AAV23197	AAV23197 Lactococc
15	16	26.7	60	AAV23214	AAV23214 Lactococc
16	16	26.7	60	AAV23204	AAV23204 Lactococc
17	16	26.7	478	AC112896	AC112896 Human adu
18	16	26.7	480	AAV23209	AAV23209 Rat seque
19	16	26.7	493	ABK62309	ABK62309 Rat seque
20	16	26.7	493	ABK62309	ABK62309 Rat seque

21	16	26.7	493	10	ADB55462	ADB55462 Toxicity-
22	16	26.7	566	13	ADB55575	ADB55575 Cotton CD
23	16	26.7	610	6	ABQ17772	ABQ17772 Oligonuc
24	16	26.7	610	6	ABQ17773	ABQ17773 Oligonuc
25	16	26.7	627	6	ABN91707	ABN91707 Staphyloc
26	16	26.7	627	13	ADS03583	ADS03583 Staphyloc
27	16	26.7	960	6	AB15530	AB15530 Arabidops
28	16	26.7	1151	8	ACA29232	ACA29232 Prokaryot
29	16	26.7	1224	10	ABX06549	ABX06549 S. pneumo
30	16	26.7	1272	13	ADR91426	ADR91426 Novel S.
31	16	26.7	1467	8	ACA21030	ACA21030 Prokaryot
32	16	26.7	1479	9	ADA31462	ADA31462 DNA encod
33	16	26.7	1549	2	AAV43003	AAV43003 Streptoco
34	16	26.7	2511	8	ACA29314	ACA29314 Prokaryot
35	16	26.7	2558	10	ADE28072	ADE28072 Corn cycl
36	16	26.7	2558	12	ADQ95183	ADQ95183 Corn cycl
37	16	26.7	2864	12	ADF50636	ADF50636 Murine DN
38	16	26.7	2866	8	ACC46187	ACC46187 Human dit
39	16	26.7	2866	12	ADK70265	ADK70265 Respirato
40	16	26.7	3320	4	AB129250	AB129250 Drosophila
41	16	26.7	3794	13	ADQ95761	ADQ95761 ECAT4 mou
42	16	26.7	3807	13	ABD33342	ABD33342 Murine ca
43	16	26.7	4547	10	ADB69145	ADB69145 C. neofo
44	16	26.7	4918	13	ADQ95768	ADQ95768 ECAT4 mou
45	16	26.7	6033	6	AA563337	AA563337 Chemical

ALIGNMENTS

RESULT 1	AAV23208	standard; DNA; 60 BP.
ID	AAV23208	standard; DNA; 60 BP.
XX	AAV23208	(first entry)
DT	28-UTR-1998	(first entry)
XX	Lactococcus lactis	constititional promoter Cp42.
DE	Lactococcus lactis	constititional promoter; optimalse; spacer;
XX	Lactococcus lactis	constititional promoter library; gene expression; ds.
KW	artificial promoter library; gene expression; ds.	
XX	Synthetic.	
OS	Lactococcus lactis.	
XX	Lactococcus lactis.	
FT	Key	Location/Qualifiers
FT	promoter	4..60
FT		/tag= a
FT		/standard_name= "Constititional promoter"
XX	WO9807846-A1.	
XX	26-FEB-1998.	
XX	25-AUG-1997;	97WO-DK00342.
XX	23-AUG-1996;	96DK-00000886.
XX	(JENS/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
XX	WPI; 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX	Claim 28; Page 57; 89pp; English.	
XX	This is a Lactococcus lactis constititional promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds

(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-36

Perfect score: 60
Sequence: 1 CATTGCTAGTATTTCTTG.....TATATTAATAAGTACTGT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : Genembl:*

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	6	BD006967
2	21	35.0	150311	10	AC112081
3	20	33.3	1386		CR848426
4	19	31.7	129989	2	AC145806
5	19	31.7	167357	2	AC120223
6	19	31.7	198566	2	EX293991
7	19	31.7	201264	2	EX950203
8	19	31.7	206709	10	AL772319
9	19	31.7	281494	2	AL103182
10	18	30.0	39716	9	HSAC000356
11	18	30.0	64670	2	AC136714_3
12	18	30.0	66749	2	AC090981
13	18	30.0	71909	2	AC105192
14	18	30.0	109355	2	AF467807
15	18	30.0	116542	5	AB075928
16	18	30.0	145797	2	AC116938
17	18	30.0	154616	2	AC149884
18	18	30.0	155420	8	AC025906
19	18	30.0	155719	2	AC117929

c 20	18	30.0	158465	9	AC068446
c 21	18	30.0	159205	9	AC027510
c 22	18	30.0	160034	10	AC131986
c 23	18	30.0	165797	2	CR847851
c 24	18	30.0	166906	5	AC146480
c 25	18	30.0	170458	9	AC146100
c 26	18	30.0	170597	5	AL929558
c 27	18	30.0	177255	9	AC136352
c 28	18	30.0	178583	2	CR788312
c 29	18	30.0	189940	5	AC147062
c 30	18	30.0	190222	5	AL935305
c 31	18	30.0	190427	2	AC119057
c 32	18	30.0	192074	10	AL840639
c 33	18	30.0	202922	2	EX901959
c 34	18	30.0	208341	2	CR387998
c 35	18	30.0	220832	2	AC126145
c 36	18	30.0	226614	2	EX957306
c 37	18	30.0	235453	10	AC133910
c 38	18	30.0	264110	2	AC122626
c 39	18	30.0	303145	8	AE017095
c 40	18	30.0	344458	2	AC131978
c 41	18	30.0	347050	3	PER929351
c 42	17	28.3	59	6	BD006949
c 43	17	28.3	60	6	BD006939
c 44	17	28.3	60	6	BD006964
c 45	17	28.3	349	6	AX246278

ALIGNMENTS

RESULT 1	BD006967	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006967	Artificial promoter libraries for selected organisms and promoters derived from such libraries.			
DEFINITION	BD006967.1	GI:18635338			
ACCESSION	BD006967	JP 2001503249-A/36.			
VERSION	JP 2001503249-A/36.				
KEYWORDS	Lactococcus lactis				
SOURCE	Lactococcus lactis				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.				
REFERENCE	1 (bases 1 to 60)				
AUTHORS	Hammer, K. and Janssen, P.R.				
TITLE	Artificial promoter libraries for selected organisms and promoters derived from such libraries				
JOURNAL	Patent: JP 2001503249-A 36 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
	OS Lactococcus lactis				
	PN JP 2001503249-A/36				
	PD 13-MAR-2001				
	PF 25-AUG-1997 JP 1998510287				
	PR 23-AUG-1996 DK 0886/96				
	PI KALIN HAMMER, PETER RUDAR JANSSEN				
	PC CI2N15/09, CI2N15/00				
	CC Strandedness: Double;				
	CC Topology: linear;				
	FT Key				
	Location/Qualifiers				
FEATURES	promoter (4)..(60).				
source	1..60				
	/organism="Lactococcus lactis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:1358"				
ORIGIN					
Query Match	100.0%;	Score 60;	DB 6;	Length 60;	
Best Local Similarity	100.0%;	Pred. No. 2.1e-24;			
Matches	60;	Conservative	0;	Mismatches	0;
		Indels	0;	Gaps	0;
1	CATTGCTAGTATTTCTTGACACCTGAGAGCGCGTGAATATATTAAGTACTGTT	60			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-36

Sequence: 1 CATTCTGTAAGTATTCTCTG.....TATTAATAAAGTACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	35.0	686	8	BZ206712 CH230-379
2	20	33.3	397	1	AL967249
3	20	33.3	470	5	BX749087 BX749087
4	20	33.3	652	1	AL899939 AL899939
5	20	33.3	654	1	AL899939 AL899939
6	20	33.3	654	1	AL899939 AL899939
7	20	33.3	659	1	AL854474
8	20	33.3	678	5	BX781831
9	20	33.3	680	1	AL891599
10	20	33.3	708	5	BX756696
11	20	33.3	732	5	BX773311
12	20	33.3	773	7	CR430899
13	20	33.3	805	7	CP217955
14	20	33.3	827	5	BX749130
15	20	33.3	867	5	BX778269
16	19	31.7	628	9	CE335717
17	19	31.7	817	9	BZ173236
18	19	31.7	824	9	CG065864
19	19	31.7	824	9	CG065867
20	19	31.7	960	5	BQ944516
21	19	31.7	1295	8	BZ575570
22	19	31.7	13059	3	AK029677
23	18	30.0	123	5	BQ667492
24	18	30.0	146	5	BQ125055

C 25	18	30.0	172	5	BQ667395
C 26	18	30.0	393	8	BH404621
C 27	18	30.0	414	1	AL895477
C 28	18	30.0	438	4	BM129960
C 29	18	30.0	512	5	BX749088
C 30	18	30.0	529	4	BM130163
C 31	18	30.0	532	1	AL8999348
C 32	18	30.0	532	4	BM130150
C 33	18	30.0	546	8	BH380632
C 34	18	30.0	602	4	BG613378
C 35	18	30.0	677	9	CE621041
C 36	18	30.0	747	6	CD468609
C 37	18	30.0	816	8	AZ706968
C 38	18	30.0	837	7	CK794848
C 39	18	30.0	1020	9	AG33262
C 40	18	30.0	1359	8	BZ574725
C 41	17	28.3	248	7	CF245643
C 42	17	28.3	249	6	CA854713
C 43	17	28.3	284	2	BJ688395
C 44	17	28.3	284	2	BB484185
C 45	17	28.3	324	7	CO188601

ALIGNMENTS

RESULT 1
LOCUS BZ206712
DEFINITION CH230-379A20.TVB CHORI-230 Segment 2 Rattus norvegicus genomic clone CH230-379A20, genomic survey sequence.
ACCESSION BZ206712
VERSION BZ206712.1 GI:23864764
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 686)
Zhang, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment Unpublished (1999)
Other GSSs: CH230-379A20.TVB
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

FEATURES
source
Location/Qualifiers
1..686
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SaNHsd/MCM"
/db_xref="taxon:10116"
/clone="CH230-379A20"
/sex="Female"
/cell_type="Brain"
/clone_1b="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 195.613 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-37

Perfect score: 59
Sequence: 1 CATCGGCTGTTATTCTTG.....TATAATGTTCACTACTGTT 59

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	2	AAV23209 Lactococc
2	18	30.5	60	2	AAV23212 Lactococc
3	18	30.5	5259	8	ADA70437 Rice gene
4	17	28.8	58	2	AAV23184 Lactococc
5	17	28.8	60	2	AAV23197 Lactococc
6	17	28.8	60	2	AAV23213 Lactococc
7	17	28.8	207	6	ABK98627 L. lactis
8	17	28.8	207	6	ACD13878 L. lactis
9	17	28.8	242	6	ABK98617 L. lactis
10	17	28.8	242	6	ACD13868 L. lactis
11	17	28.8	1520	4	AAH29876 C albican
12	17	28.8	7431	2	AAQ30148 Attenuate
13	17	28.8	7432	2	AAQ32965 True type
14	17	28.8	8905	2	AAK20263 Borrelia
15	17	28.8	9566	6	ABK98634 Vector ps
16	17	28.8	9566	9	ACD13885 L. lactis
17	17	28.8	110000	6	ABA90521 Genomic s
18	17	28.8	194534	12	ADA97481 Human can
19	16	27.1	20	10	ABZ89000 Human oli
20	16	27.1	20	11	ABD55230 A1051839-

21	16	27.1	60	2	AAV23188 Lactococc
22	16	27.1	60	2	AAV23214 Lactococc
23	16	27.1	60	2	AAV23178 Lactococc
24	16	27.1	60	2	AAV23205 Lactococc
25	16	27.1	115	6	ABK98626 L. lactis
26	16	27.1	115	6	ACD13877 L. lactis
27	16	27.1	150	6	ABK98616 L. lactis
28	16	27.1	150	9	ACD13867 L. lactis
29	16	27.1	265	2	AAQ23880 PKH1820
30	16	27.1	265	2	AAQ46205 Promoter
31	16	27.1	265	2	AAK31875 Promoter
32	16	27.1	473	6	ABK64666 Human nuc
33	16	27.1	473	11	ABD25170 A1051839
34	16	27.1	498	12	ADP93018 Cotton ex
35	16	27.1	544	9	ACH39683 Human foe
36	16	27.1	591	6	ABN65681 Human can
37	16	27.1	629	4	AAK89514 Human dig
38	16	27.1	629	4	AAK89515 Human dig
39	16	27.1	629	4	AAK89515 Leukaemia
40	16	27.1	1224	10	ADF82234 Full leng
41	16	27.1	3696	13	ADK06604 Full leng
42	16	27.1	3838	4	AB102605 Drosophila
43	16	27.1	4215	13	ADK07845 Full leng
44	16	27.1	4295	4	AB102603 Drosophila
45	16	27.1	4533	12	ADN14346 Human qui

ALIGNMENTS

RESULT 1	AAV23209	standard, DNA; 59 BP.
ID	AAV23209	standard, DNA; 59 BP.
XX	AAV23209;	
XX	26-JUL-1998	(first entry)
DT	Lactococcus lactis	constitutional promoter Cp44.
XX	Lactococcus lactis	constitutional promoter; optimise; spacer;
DE	Lactococcus lactis	constitutional promoter; optimise; spacer;
XX	Lactococcus lactis	constitutional promoter; optimise; spacer;
KW	artificial promoter library; gene expression; ds.	
XX	Synthetic.	
OS	Lactococcus lactis.	
XX	Lactococcus lactis.	
XX	Key	Location/Qualifiers
FT	promoter	4..59
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
XX	W09807846-A1.	
XX	26-FEB-1998.	
XX	25-AUG-1997;	97MO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
XX	(JENS/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
XX	WPI, 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX	Claim 28, Page 58; 89pp; English.	
XX	This is a Lactococcus lactis constitutional promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 504.339 Seconds

(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657b-37

Perfect score: 59
Sequence: 1 CATCGGAGTACTTATCTCTG.....TATATAGTTCAGTACTGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Genbml:.*
1: gb_ba:.*
2: gb_hng:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	6	BD006968
2	20	33.9	581	11	HS1002F8T
3	20	33.9	148042	2	AC023227
4	20	33.9	155470	9	HSDJ53F4
5	19	32.2	907	8	AY254289
6	19	32.2	112638	9	HSJ329L24
7	19	32.2	156825	9	AC026726
8	19	32.2	195771	10	AC108819
9	19	32.2	196566	10	AC118608
10	19	32.2	199760	2	AC113962
11	18	30.5	60	6	BD006971
12	18	30.5	1829	10	BC075654
13	18	30.5	5139	8	AY374515
14	18	30.5	5259	6	AX653890
15	18	30.5	5307	8	AY374517
16	18	30.5	5348	8	AY374516
17	18	30.5	38892	9	AL139145
18	18	30.5	73260	5	AL606809
19	18	30.5	114089	10	AL807770

20	18	30.5	136903	5	CR352245	CR352245 Zebrafish
21	18	30.5	139872	8	AC091811	AC091811 Oryza sat
22	18	30.5	143903	5	BX247887	BX247887 Zebrafish
23	18	30.5	147230	2	CR847998	CR847998 Danio rer
24	18	30.5	147699	9	AC006500	AC006500 Homo sapi
25	18	30.5	149380	8	AC087852	AC087852 Oryza sat
26	18	30.5	151144	5	BX323575	BX323575 Zebrafish
27	18	30.5	151619	2	CR848024	CR848024 Danio rer
28	18	30.5	152121	2	BX255868	BX255868 Danio rer
29	18	30.5	157985	9	CNS057BM	AL163853 Human chr
30	18	30.5	159698	5	AL808019	AL808019 Zebrafish
31	18	30.5	160925	8	AP004332	AP004332 Oryza sat
32	18	30.5	170325	2	CR788305	CR788305 Danio rer
33	18	30.5	171113	5	BX682531	BX682531 Zebrafish
34	18	30.5	173877	10	AC099624	AC099624 Mus muscu
35	18	30.5	178502	10	AC120397	AC120397 Mus muscu
36	18	30.5	182126	2	CR381684	CR381684 Danio rer
37	18	30.5	182619	10	AC127591	AC127591 Mus muscu
38	18	30.5	185460	5	BX537259	BX537259 Zebrafish
39	18	30.5	188746	5	AL929594	AL929594 Zebrafish
40	18	30.5	189752	2	CR381635	CR381635 Danio rer
41	18	30.5	191535	2	BX511071	BX511071 Danio rer
42	18	30.5	192213	2	CR847542	CR847542 Danio rer
43	18	30.5	193930	9	AC006296	AC006296 Homo sapi
44	18	30.5	206733	2	AC110553	AC110553 Mus muscu
45	18	30.5	215849	2	AC007939	AC007939 Homo sapi

ALIGNMENTS

RESULT 1
BD006968
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters derived from such libraries.

ACCESSION
BD006968.1 GI:18635339
VERSION
JP 2001503249-A/37.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE
1 (bases 1 to 59)
AUTHORS
Hammer, K. and Janssen, P. R.
TITLE
Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL
Patent: JP 2001503249-A 37 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT
OS Lactococcus lactis
PN JP 2001503249-A/37
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96

FEATURES
source
ORIGIN
Query Match 100.0%; Score 59; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 8.6e-24;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATCGGAGTACTTATCTGACATTAAGTAGAGCTGATATATAGTTCAGTACTGTT 59

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using 'sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1430.81 Seconds

(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-37

Perfect score: 59

Sequence: 1 CATCGGCTAGTTATTCTTG.....TATATAGTTCAGTACTGTT 59

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 1903134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18	30.5	507 8	A0704541 HS_5487_B
2	18	30.5	597 8	A0256211 LOCUS cor
3	18	30.5	646 8	A0510836 nbd0048L
4	18	30.5	654 9	AG021073 Oryza sat
5	18	30.5	656 7	CO044252 UT-M-GMO-
6	18	30.5	680 6	CB837323 VVE087A12
7	18	30.5	683 6	CB837576 VVE090A07
8	18	30.5	684 6	CB837066 VVE083B06
9	18	30.5	684 6	CB837281 VVE084F01
10	18	30.5	684 6	CB837281 VVE086D09
11	18	30.5	684 6	CB837281 VVE086D09
12	18	30.5	685 6	CB837281 VVE086D09
13	18	30.5	685 6	CB837281 VVE086D09
14	18	30.5	686 6	CB837281 VVE086D09
15	18	30.5	687 6	CB837281 VVE086D09
16	18	30.5	688 6	CB837281 VVE086D09
17	18	30.5	693 6	CB837281 VVE086D09
18	18	30.5	693 6	CB837281 VVE086D09
19	18	30.5	693 6	CB837281 VVE086D09
20	18	30.5	693 6	CB837281 VVE086D09
21	18	30.5	694 6	CB837281 VVE086D09
22	18	30.5	694 6	CB837281 VVE086D09
23	18	30.5	694 6	CB837281 VVE086D09
24	18	30.5	697 6	CB837281 VVE086D09

C 25	18	30.5	697 6	CB837121 VVE084A05
C 26	18	30.5	698 6	CB834559 VVE011B04
C 27	18	30.5	701 6	CB836797 VVE078E12
C 28	18	30.5	710 6	CB834847 VVE014F09
C 29	18	30.5	710 6	CB837237 VVE085G08
C 30	18	30.5	710 6	CB837237 VVE085G08
C 31	18	30.5	711 6	CB834572 VVE011C08
C 32	18	30.5	711 6	CB834572 VVE011C08
C 33	18	30.5	711 6	CB834572 VVE011C08
C 34	18	30.5	711 6	CB834572 VVE011C08
C 35	18	30.5	711 6	CB834572 VVE011C08
C 36	18	30.5	717 6	CB837631 VVE080H11
C 37	18	30.5	722 6	CB837748 VVE082H05
C 38	18	30.5	726 6	CB836532 VVE074A10
C 39	18	30.5	727 9	BX161687 VVE074A10
C 40	18	30.5	733 6	CB837108 VVE083G10
C 41	18	30.5	737 6	CB837257 VVE086A12
C 42	18	30.5	739 6	CB835540 VVE056B01
C 43	18	30.5	739 8	B2062012 1Kf67B07.
C 44	18	30.5	745 8	B2061945 1Kf67B07.
C 45	18	30.5	789 8	BH694879 BOM05567P

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

A0704541
HS_5487_B1 A08 SPEE RPT-11 Human Male BAC library Homo sapiens
genomic clone Plate=1063 Col=15 Row=B, genomic survey sequence.
A0704541
GI:5413967
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 507)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPT-11. For BAC
library availability, please contact Pletier de Jong
(pletier@dejong.med.buffalo.edu). Clones may be purchased from
BACpac Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1063 row: B column: 15
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 507.
Location/Qualifiers
1..507
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1063 Col=15 Row=B"
/sex="male"
/clone_lib="RPT-11 Human Male BAC library"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 / Search time 195.613 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-38

Perfect score: 59
Sequence: 1 CATGGGGAGAGTTATCTTCTTGTG.....GTATACATACAGATATGTT 59

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn19808:*\n2: geneseqn19908:*\n3: geneseqn20008:*\n4: geneseqn2001a:*\n5: geneseqn2001b:*\n6: geneseqn2002a:*\n7: geneseqn2002b:*\n8: geneseqn2003a:*\n9: geneseqn2003b:*\n10: geneseqn2003c:*\n11: geneseqn2003d:*\n12: geneseqn2004a:*\n13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	59	AAV23210	AAV23210 Lactococc
2	2000	33.9	8	ADA72341	ADA72341 Rice gene
3	60	32.2	60	AAV23199	AAV23199 Lactococc
4	18	30.5	60	AAV23211	AAV23211 Lactococc
5	18	30.5	60	AAV23200	AAV23200 Lactococc
6	18	30.5	12	AD059380	AD059380 Human can
7	17	28.8	60	AAV23177	AAV23177 Lactococc
8	17	28.8	60	AAV23191	AAV23191 Lactococc
9	17	28.8	60	AAV23182	AAV23182 Lactococc
10	17	28.8	417	AAI13124	AAI13124 Probe #12
11	17	28.8	417	AAI13124	AAI13124 Probe #12
12	17	28.8	417	AAI13124	AAI13124 Probe #12
13	17	28.8	417	AAI13124	AAI13124 Probe #12
14	17	28.8	417	AAI13124	AAI13124 Probe #12
15	17	28.8	417	AAI13124	AAI13124 Probe #12
16	17	28.8	417	AAI13124	AAI13124 Probe #12
17	17	28.8	417	AAI13124	AAI13124 Probe #12
18	17	28.8	417	AAI13124	AAI13124 Probe #12
19	17	28.8	417	AAI13124	AAI13124 Probe #12
20	17	28.8	417	AAI13124	AAI13124 Probe #12

C	21	17	28.8	417	5	AAI03027	AAI03027 Probe #30
C	22	17	28.8	417	6	AB03063	AB03063 Human gen
C	23	17	28.8	417	6	AB015617	AB015617 Human gen
C	24	17	28.8	419	8	ABX55241	ABX55241 Bovine ES
	25	17	28.8	1969	4	AAI14105	AAI14105 Probe #40
	26	17	28.8	1969	4	AB055830	AB055830 Human foe
	27	17	28.8	1969	4	AAI35486	AAI35486 Probe #41
	28	17	28.8	1969	4	AB045341	AB045341 Human bre
	29	17	28.8	1969	4	AB025506	AB025506 Probe #39
	30	17	28.8	1969	4	AAK29533	AAK29533 Human don
	31	17	28.8	1969	4	AAK04048	AAK04048 Human bra
	32	17	28.8	1969	4	AB029157	AB029157 Human liv
	33	17	28.8	1969	5	AAI03958	AAI03958 Probe #39
	34	17	28.8	1969	6	AB04084	AB04084 Human gen
C	35	17	28.8	48829	11	ACN44868	ACN44868 Mouse gen
C	36	17	28.8	110000	6	AB03041_13	Continuation (14 O
	37	16	27.1	58	2	AAV23186	AAV23186 Lactococc
	38	16	27.1	59	2	AAV23190	AAV23190 Lactococc
	39	16	27.1	59	2	AAV23185	AAV23185 Lactococc
	40	16	27.1	59	2	AAV23193	AAV23193 Lactococc
	41	16	27.1	59	2	AAV23206	AAV23206 Lactococc
	42	16	27.1	60	2	AAV23194	AAV23194 Lactococc
	43	16	27.1	60	2	AAV23188	AAV23188 Lactococc
	44	16	27.1	60	2	AAV23189	AAV23189 Lactococc
	45	16	27.1	60	2	AAV23198	AAV23198 Lactococc

ALIGNMENTS

RESULT 1	AAV23210	standard; DNA; 59 BP.
ID	AAV23210	
XX	AAV23210;	
AC	AAV23210;	
XX	AAV23210;	
DT	28-JUL-1998	(first entry)
XX	28-JUL-1998	
DE	Lactococcus lactis	constitutional promoter Cps.
XX	Lactococcus lactis	constitutional promoter; optimise; spacer.
KW	lactococcus lactis; constitutional promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; de.	
OS	Synthetic.	
OS	Lactococcus lactis.	
XX	Lactococcus lactis.	
PH	Key	Location/Qualifiers
FT	promoter	4..59
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
XX	W09807846-A1.	
XX	26-FEB-1998.	
PD	26-FEB-1998.	
XX	25-AUG-1997;	97MO-DK000342.
PF	25-AUG-1997;	
XX	23-AUG-1996;	96DX-00000886.
PR	23-AUG-1996;	
XX	(JENS/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
PI	Jensen PR, Hammer K;	
XX	WPI; 1998-179062/16.	
DR	WPI; 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX	Claim 28; Page 58; 89pp; English.	
PS	This is a Lactococcus lactis constitutional promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	
CC		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 504.339 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-38

Perfect score: 59
Sequence: 1 CATGGGGAGATTATTTCTTG.....GTATCTACATAGATGTT 59

Scoring table: Oligo NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmb1:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	59	6	BD006969
2	20	33.9	1470	8	AK119468
3	20	33.9	2000	6	AX655796
4	20	33.9	3378	8	AK099595
5	20	33.9	101901	8	AP003377
6	20	33.9	137770	8	AP003371
7	20	33.9	151303	8	AP003793
8	19	32.2	60	6	BD006958
9	19	32.2	86422	9	AC105362
10	19	32.2	228016	2	AC146674
11	19	32.2	228049	2	AC095210
12	19	32.2	239786	2	AC119498
13	19	32.2	270679	2	AC118185
14	19	32.2	275980	2	AC111446
15	18	30.5	60	6	BD006959
16	18	30.5	109588	6	BD006970
17	18	30.5	117522	9	AC004743
18	18	30.5	154772	9	AL138764
19	18	30.5	154772	9	AL138764

20	18	30.5	161920	9	AC025169
21	18	30.5	214786	9	AC007610
22	18	30.5	233058	5	BX640463
23	17	28.8	60	6	BD006936
24	17	28.8	60	6	BD006941
25	17	28.8	60	6	BD006950
26	17	28.8	148	9	HSU32637
27	17	28.8	187	9	HSU32639
28	17	28.8	351	9	HSU32642
29	17	28.8	417	6	CQ052198
30	17	28.8	417	6	CQ057215
31	17	28.8	417	6	CQ067257
32	17	28.8	417	6	CQ076493
33	17	28.8	417	6	CQ094302
34	17	28.8	417	6	CQ107477
35	17	28.8	417	6	CQ133086
36	17	28.8	417	6	CQ146143
37	17	28.8	417	6	CQ171657
38	17	28.8	417	6	CQ181584
39	17	28.8	417	6	CQ200788
40	17	28.8	417	6	CQ205941
41	17	28.8	417	6	CQ291949
42	17	28.8	417	6	CQ304503
43	17	28.8	515	8	AY130795
44	17	28.8	515	8	AY130796
45	17	28.8	584	11	BV059665

ALIGNMENTS

RESULT 1
BD006969
LOCUS
DEFINITION Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION BD006969.1 GI:18635340
VERSION JP 2001503249-A/38.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

REFERENCE
AUTHORS Hammer, K. and Janssen, P.R.
TITLE Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL Patent: JP 2001503249-A 38 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Lactococcus lactis
PN JP 2001503249-A/38
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC Cl2N15/09, Cl2N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (4)..(59).
Location/Qualifiers
1..59
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 59; DB 6; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGGGGAGATTATTTCTTGACATCATCTGAGCCCTGATACATAGATGTT 59

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1430.81 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657b-38

Perfect score: 59

Sequence: 1 CATGGGGAGATTATCTTG.....GTATACATAGATGTT 59

Scoring table: OLIGO_NUC

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	33.9	323	8	A0456217 HS 5076 B
2	20	33.9	805	8	CB677677 OSJNBe10
3	19	32.2	295	9	AG196813 Pan trogl
4	18	30.5	240	1	AJ483008
5	18	30.5	300	1	AJ481220
6	18	30.5	360	1	AJ483010
7	18	30.5	420	1	AJ483004
8	18	30.5	420	1	AJ483006
9	18	30.5	427	5	B0882237
10	18	30.5	441	6	BY572200
11	18	30.5	480	1	AJ483005
12	18	30.5	480	1	AJ483007
13	18	30.5	540	1	AJ483007
14	18	30.5	540	1	AJ483009
15	18	30.5	643	5	BQ764357
16	18	30.5	652	9	DR15167
17	18	28.8	927	9	CG703804
18	17	28.8	300	1	AJ484038
19	17	28.8	327	2	AM356517
20	17	28.8	344	2	AM457877
21	17	28.8	347	2	AM457877
22	17	28.8	358	6	CF029617
23	17	28.8	368	6	CA935425
24	17	28.8	382	4	BMS68009

c 25	17	28.8	420	2	AM156579
c 26	17	28.8	436	6	CF030799
c 27	17	28.8	436	6	CF030858
c 28	17	28.8	436	6	CF030875
c 29	17	28.8	436	6	CF031004
c 30	17	28.8	436	6	CF031256
c 31	17	28.8	437	6	CF030997
c 32	17	28.8	438	6	CF029631
c 33	17	28.8	438	6	CF030356
c 34	17	28.8	438	6	CF030598
c 35	17	28.8	456	6	CF058870
c 36	17	28.8	457	6	CF030597
c 37	17	28.8	458	1	AT855241
c 38	17	28.8	458	7	CF438820
c 39	17	28.8	460	6	CF030874
c 40	17	28.8	467	6	CD988414
c 41	17	28.8	473	6	CF030752
c 42	17	28.8	475	5	B0161553
c 43	17	28.8	479	6	CF030599
c 44	17	28.8	484	6	CF031602
c 45	17	28.8	485	6	CF030870

ALIGNMENTS

RESULT 1
A0456217
LOCUS
DEFINITION
A0456217 HS 5076 B1 C10 SP6E RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=652 Col=19 Row=F, genomic survey sequence.
ACCESSION
A0456217
VERSION
A0456217.1 GI:4591552
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
A0456217
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.

TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
PUBMED
10449764
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACpac Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 652 row: F column: 19
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 323.
Location/Qualifiers
1. 323

FEATURES
source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=652 Col=19 Row=F"
/sex="male"
/clone_1lb="RPCI-11 Human Male BAC Library"
/notes="vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.922 Seconds

(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-39

Sequence: 1 CATGTGGAGTTATCTCTG.....TATTAATAGTACTGTT 60

Scoring table: OLIGO_NUC

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23211
2	20	33.3	59	2	AAV23196
3	20	33.3	60	2	AAV23182
4	18	30.0	59	2	AAV23210
5	18	30.0	60	2	AAV23199
6	17	28.3	59	2	AAV23190
7	17	28.3	60	2	AAV23177
8	17	28.3	60	2	AAV23191
9	17	28.3	60	2	AAV23197
10	17	28.3	60	2	AAV23200
11	17	28.3	1039	3	AAV23200
12	17	28.3	1039	3	AAV23200
13	17	28.3	1039	3	AAV23200
14	17	28.3	2864	12	ADP18784
15	16	26.7	59	2	AAV23185
16	16	26.7	60	2	AAV23208
17	16	26.7	60	2	AAV23189
18	16	26.7	60	2	AAV23214
19	16	26.7	60	2	AAV23198
20	16	26.7	60	2	AAV23180

21	16	26.7	60	2	AAV23205	AAV23205 Lactococc
22	16	26.7	265	2	AAQ23880	AAQ23880 PKT1820
23	16	26.7	265	2	AAQ46205	AAQ46205 Promoter
24	16	26.7	265	2	AAQ31875	AAQ31875 Promoter
25	16	26.7	296	3	AAQ31364	AAQ31364 Human sec
26	16	26.7	307	8	ABX43872	ABX43872 Bovine ES
27	16	26.7	391	4	AAI10483	AAI10483 Human bre
28	16	26.7	569	4	AAI35488	AAI35488 Human mus
29	16	26.7	569	8	ABX58476	ABX58476 CDNA enco
30	16	26.7	569	12	ADJ28203	ADJ28203 Human mus
31	16	26.7	651	3	AAI13601	AAI13601 Aspergill
32	16	26.7	821	4	AAI17559	AAI17559 Human bre
33	16	26.7	860	11	ACN080681	ACN080681 Breast ca
34	16	26.7	1639	4	AAQ08090	AAQ08090 Beetele cd
35	16	26.7	1639	6	ABQ81817	ABQ81817 Photuris
36	16	26.7	1639	12	ADM31222	ADM31222 Mutant lu
37	16	26.7	1639	13	ADQ91605	ADQ91605 Luciferas
38	16	26.7	2308	11	ACN89898	ACN89898 Breast ca
39	16	26.7	38239	12	ADQ97626	ADQ97626 Mouse can
40	16	26.7	54701	11	ACN44478	ACN44478 Human gen
41	16	26.7	66312	11	ACN44550	ACN44550 Human gen
42	16	26.7	93329	13	ABD33597	ABD33597 Murine ca
43	16	26.7	110000	6	ABA90521_02	ABA90521_02 Continuation (3 of
44	16	26.7	202001	6	AB552506	AB552506 Human tra
45	16	26.7	202001	10	ADG46742	ADG46742 Human tra

ALIGNMENTS

RESULT 1	AAV23211	standard; DNA; 60 BP.
ID	AAV23211	
XX	AAV23211;	
AC	28-JUL-1998	(first entry)
XX		
DT	Lactococcus lactis	constitutional promoter Cp6.
XX		
DE	Lactococcus lactis	constitutional promoter; optimise; spacer;
XX		
KW	artificial promoter library; gene expression; ds.	
XX		
OS	Synthetic.	
OS	Lactococcus lactis.	
XX		
FT	Key	Location/Qualifiers
FT	Promoter	4..60
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
PN	WO9807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PP	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENSEN) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
PS	Claim 28; Page 59; 89pp; English.	
XX		
CC	This is a Lactococcus lactis constitutional promoter sequence used in the	
	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657b-39

Perfect score: 60
Sequence: 1 CAGTGGAGATTCTTCTTG.....TATTAATGAGTACTGTT 60

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:**

1: gb_Pa:*
2: gb_Htg:*
3: gb_In:*
4: gb_Om:*
5: gb_Ov:*
6: gb_Pat:*
7: gb_Pn:*
8: gb_Pl:*
9: gb_Pr:*
10: gb_Ro:*
11: gb_Sts:*
12: gb_Sy:*
13: gb_Un:*
14: gb_Vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006970	Artificia
2	20	33.3	59	BD006955	Artificia
3	20	33.3	60	BD006941	Artificia
4	19	31.7	55957	AC110155	Artificia
5	19	31.7	84553	AC123538	Mus muscu
6	19	31.7	110000	AC151267	Mus muscu
7	19	31.7	157999	AC117088	Rattus no
8	19	31.7	165281	CR388133	Danio rer
9	19	31.7	231150	AC114195	Rattus no
10	19	31.7	277191	AC109077	Rattus no
11	18	30.0	59	BD006969	Artificia
12	18	30.0	60	BD006958	Artificia
13	18	30.0	798	BY050056	S21286965
14	18	30.0	1107	CNS06K22	T3 end of
15	18	30.0	139015	AC087581	Homo sapi
16	18	30.0	139032	AC010656	Homo sapi
17	18	30.0	156957	AP001977	Homo sapi
18	18	30.0	168365	AC131653	Mus muscu
19	18	30.0	172346	AC016349	Homo sapi

20	18	30.0	179737	2	AC121552	Mus muscu
21	18	30.0	192174	2	AC128248	Rattus no
22	18	30.0	240832	2	AC137395	Rattus no
23	18	30.0	244444	2	AC099182	Rattus no
24	18	30.0	261240	2	AC118604	Mus muscu
25	18	30.0	267937	2	AC108665	Rattus no
26	18	30.0	272027	2	AC095229	Rattus no
27	18	30.0	280815	2	AC107362	Rattus no
28	18	30.0	301629	2	AC103342	Rattus no
29	18	30.0	314656	2	AC099366	Rattus no
30	18	30.0	343558	2	AC117032	Rattus no
31	18	30.0	344615	2	AC125938	Rattus no
32	17	28.3	59	6	BD006949	Artificia
33	17	28.3	60	6	BD006936	Artificia
34	17	28.3	60	6	BD006950	Artificia
35	17	28.3	60	6	BD006956	Artificia
36	17	28.3	60	6	BD006959	Artificia
37	17	28.3	106	11	CR378556	Arabidops
38	17	28.3	931	9	BC033933	Homo sapi
39	17	28.3	1039	6	AR306584	Sequence
40	17	28.3	1039	6	AX061658	Sequence
41	17	28.3	1063	3	AY301327	Antheraea
42	17	28.3	12850	1	AY545992	Baccharich
43	17	28.3	24948	1	AY762939	Klebsiell
44	17	28.3	30321	8	AC138003	Oryza sat
45	17	28.3	33415	1	AB117611	Klebsiell

ALIGNMENTS

RESULT 1
LOCUS BD006970
DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION BD006970.1 GI:18635341
VERSION JP 2001503249-A/39.
KEYWORDS Lactococcus lactis
SOURCE Lactococcus lactis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE 1 (bases 1 to 60)
AUTHORS Hammer,K. and Janssen,P.R.
TITLE Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL Patent: JP 2001503249-A 39 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT OS Lactococcus lactis
PN JP 2001503249-A/39
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PI 23-AUG-1996 DK 0886/96
PL KALIN HAMMER, PETER RUDAR JANSSEN
PC G12N15/09, G12N15/00
CC Strandness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (4). (60).

FEATURES
source Location/Qualifiers
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.6e-24;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGGAGATTCTTCTTGACACAGATTTCCGATGATATTAATGAGTACTGTT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds

(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-39

Perfect score: 1 CATGTGGAGATTATTCTTG.....TATATATACGTGACTGTT 60

Sequence: 1 CATGTGGAGATTATTCTTG.....TATATATACGTGACTGTT 60

Scoring table: OLIGO_NDC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	33.3	579	9	CE530345
C 2	19	31.7	710	9	CR217070
C 3	19	31.7	912	9	CR187572
C 4	18	30.0	197	6	CB062220
C 5	18	30.0	289	8	BH296821
C 6	18	30.0	443	8	AZ182039
C 7	18	30.0	576	8	AZ159092
C 8	18	30.0	698	8	BH942434
C 9	18	30.0	699	8	BZ173504
C 10	18	30.0	699	8	BZ173504
C 11	18	30.0	768	7	CN188268
C 12	18	30.0	852	6	CB292840
C 13	18	30.0	927	6	CG703804
C 14	17	28.3	358	6	CF029617
C 15	17	28.3	361	4	BI019191
C 16	17	28.3	365	4	AI080135
C 17	17	28.3	405	8	AO597540
C 18	17	28.3	425	1	AI571998
C 19	17	28.3	430	5	EX119079
C 20	17	28.3	436	6	CF030799
C 21	17	28.3	436	6	CF030858
C 22	17	28.3	436	6	CF030875
C 23	17	28.3	436	6	CF031004
C 24	17	28.3	436	6	CF031256

C 25	17	28.3	437	6	CF030997	CF030997
C 26	17	28.3	438	6	CF029631	CF029631
C 27	17	28.3	438	6	CF030356	CF030356
C 28	17	28.3	438	6	CF030598	CF030598
C 29	17	28.3	440	4	BG192478	BG192478
C 30	17	28.3	440	8	AO661689	AO661689
C 31	17	28.3	456	6	CF058870	CF058870
C 32	17	28.3	457	6	CF050597	CF050597
C 33	17	28.3	458	1	AI855241	AI855241
C 34	17	28.3	460	6	CF030874	CF030874
C 35	17	28.3	467	6	CD988414	CD988414
C 36	17	28.3	473	6	CF030752	CF030752
C 37	17	28.3	474	1	AI634177	AI634177
C 38	17	28.3	479	6	CF030599	CF030599
C 39	17	28.3	480	4	BM344151	BM344151
C 40	17	28.3	484	6	CF031602	CF031602
C 41	17	28.3	485	6	CF030870	CF030870
C 42	17	28.3	486	6	CF031026	CF031026
C 43	17	28.3	486	6	CF031591	CF031591
C 44	17	28.3	487	1	AI095614	AI095614
C 45	17	28.3	500	4	BM059694	BM059694

ALIGNMENTS

RESULT 1
CE530345/c
LOCUS
DEFINITION
tigr-gss-dog-17000365867028 Dog Library Canis familiaris genomic,
genomic survey sequence.
CE530345
VERSION
CE530345.1 GI:36847126
KEYWORDS
GSS.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
REFERENCE
Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 579)
AUTHORS
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE
The dog genome: survey sequencing and comparative analysis
JOURNAL
Science 301 (5641), 1898-1903 (2003)
MEDLINE
22875432
PUBMED
14512627

COMMENT

Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: kirkness@tigr.org
Class: shotgun.
Location/Qualifiers
1. 579
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Bcfxi; Libraries were prepared from
peripheral blood"

FEATURES

ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
36 GATGATATATATACGTAGTA 55
493 GATGATATATATACGTAGTA 474

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657b-40

Perfect score: 60

Sequence: 1 TATGCGGTAGTTATCTCTG.....TATATGGGTCTAGATTAGG 60

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: GenBank1:

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pa:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006971	BD006971 Artificial
2	19	31.7	60	BD006947	BD006947 Artificial
3	19	31.7	206016	AC133808	AC133808 Rattus no
4	19	31.7	262424	AC103010	AC103010 Rattus no
5	18	30.0	59	BD006968	BD006968 Artificial
6	18	30.0	21462	AC006339	AC006339 Homo sapi
7	18	30.0	169358	AL158165	AL158165 Human DNA
8	18	30.0	180301	AC146083	AC146083 Pan trogl
9	17	28.3	58	BD006943	BD006943 Artificial
10	17	28.3	59	BD006952	BD006952 Artificial
11	17	28.3	60	BD006956	BD006956 Artificial
12	17	28.3	1971	PT4GP2A	PT4GP2A Bacterioph
13	17	28.3	2859	MYT4501R	MYT4501R Sequence
14	17	28.3	15320	CO581234	CO581234 Drosophi
15	17	28.3	48042	AC020217	AC020217 Drosophi
16	17	28.3	129908	AR265715	AR265715 Sequence
17	17	28.3	129908	AX059140	AX059140 Sequence
18	17	28.3	133659	AC092094	AC092094 Mus muscu
19	17	28.3	152719	AC012545	AC012545 Homo sapi

20	17	28.3	153070	9	AL355361	AL355361 Human DNA
21	17	28.3	154513	2	AC145047	AC145047 Sus scrofa
22	17	28.3	161775	9	AC099512	AC099512 Homo sapi
23	17	28.3	162591	9	AL391839	AL391839 Human DNA
24	17	28.3	164461	9	AC083806	AC083806 Homo sapi
25	17	28.3	164567	9	AC021660	AC021660 Homo sapi
26	17	28.3	165362	10	AL844862	AL844862 Mouse DNA
27	17	28.3	166720	10	AC125099	AC125099 Mus muscu
28	17	28.3	168903	7	AF158101	AF158101 Enterobac
29	17	28.3	169705	3	AC104704	AC104704 Drosophi
30	17	28.3	171907	2	AC116538	AC116538 Drosophi
31	17	28.3	172752	2	AC116496	AC116496 Mus muscu
32	17	28.3	176194	9	AL691486	AL691486 Human DNA
33	17	28.3	176697	2	AC130526	AC130526 Rattus no
34	17	28.3	177987	9	AC107052	AC107052 Homo sapi
35	17	28.3	180689	2	AC124913	AC124913 Sus scrofa
36	17	28.3	182918	2	AC104932	AC104932 Mus muscu
37	17	28.3	183193	2	BX936312	BX936312 Dario rer
38	17	28.3	183371	9	HS625H18	HS625H18 Human DNA
39	17	28.3	185371	9	AC018622	AC018622 Homo sapi
40	17	28.3	187574	2	CR383675	CR383675 Dario rer
41	17	28.3	190425	2	AC144593	AC144593 Mus muscu
42	17	28.3	193046	2	AC062034	AC062034 Homo sapi
43	17	28.3	193902	2	AL591112	AL591112 Homo sapi
44	17	28.3	222577	2	AC114170	AC114170 Rattus no
45	17	28.3	222614	2	AC123961	AC123961 Rattus no

ALIGNMENTS

RESULT 1
BD006971
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006971.1 GI:18635342
ACCESSION
VERSION
UP 2001503249-A/40.
KEYWORDS
Lactococcus lactis
SOURCE
ORGANISM
Lactococcus lactis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.

REFERENCE
1 (bases 1 to 60)

AUTHORS
Hammer, K. and Janssen, P.R.

TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 40 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT
OS Lactococcus lactis
PN UP 2001503249-A/40
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PF 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN

FEATURES
source
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TATGCGGTAGTTATCTCTGACATGACGACAGGTGTGTATATGGCTCTAGATTAGG 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 21:47:22 ; Search time 512.887 Seconds
(without alignments)
5668.525 Million cell updates/sec

Title: US-09-242-657B-41

Sequence: 1 CATCTTGTAGTTATCTTG.....TATATAGGTGAGTACTGTT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Geneml:.*
1: gb_ba:.*
2: gb_hgt:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_scs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006972	BD006972 Artificia
2	23	38.3	60	BD006959	BD006959 Homo sapi
3	19	31.7	88280	AC138407	AC138407 Homo sapi
4	19	31.7	121280	AC087045	AC087045 Homo sapi
5	19	31.7	142331	AF307157	AF307157 Homo sapi
6	19	31.7	159032	AC083843	AC083843 Homo sapi
7	19	31.7	160925	AP004332	AP004332 Oryza sat
8	19	31.7	172826	AL671917	AL671917 Mouse DNA
9	19	31.7	198014	AC130026	AC130026 Rattus no
10	19	31.7	241640	AC142065	AC142065 Rattus no
11	19	31.7	259682	AC111230	AC111230 Rattus no
12	18	30.0	60	BD006964	BD006964 Artificia
13	18	30.0	67	AR055663	AR055663 Sequence
14	18	30.0	1100	1.LITRRM	X75364 L.lactis ge
15	18	30.0	1100	6.AB055662	AR055662 Sequence
16	18	30.0	12737	1.AB006241	AB006241 Lactococc
17	18	30.0	44274	9.AL954338	AL954338 Human DNA
18	18	30.0	73840	8.AB007644	AB007644 Arabidops
19	18	30.0	82518	2.AC092077	AC092077 Oryza sat

20	18	30.0	87574	9.AC084010	AC084010 Homo sapi
21	18	30.0	110000	2.AC111295_2	Continuation (3 of
22	18	30.0	123825	9.HS782123	AL035416 Human DNA
23	18	30.0	124872	10.AC132094	AC132094 Mus muscu
24	18	30.0	142224	2.AC096535	AC096535 Homo sapi
25	18	30.0	145155	8.AC091233	AC091233 Oryza sat
26	18	30.0	151924	9.AC098648	AC098648 Homo sapi
27	18	30.0	156791	10.AC121504	AC121504 Mus muscu
28	18	30.0	164006	5.BX321505	BX321505 Zebrafish
29	18	30.0	165220	5.BX300080	BX300080 Zebrafish
30	18	30.0	180223	2.AC068979	AC068979 Homo sapi
31	18	30.0	186589	2.AC141218	AC141218 Rattus no
32	18	30.0	192940	5.AL845512	AL845512 Zebrafish
33	18	30.0	203639	10.AC123876	AC123876 Mus muscu
34	18	30.0	205597	2.AC142481	AC142481 Rattus no
35	18	30.0	215731	2.AC136554	AC136554 Rattus no
36	18	30.0	239569	2.AC135812	AC135812 Rattus no
37	18	30.0	248894	2.AC098353	AC098353 Rattus no
38	18	30.0	254007	2.AC094413	AC094413 Rattus no
39	18	30.0	254483	2.AC097072	AC097072 Rattus no
40	18	30.0	256473	2.AC109689	AC109689 Rattus no
41	18	30.0	260045	2.AC116843	AC116843 Mus muscu
42	18	30.0	266999	2.AC121640	AC121640 Rattus no
43	17	28.3	59	6.BD006968	BD006968 Artificia
44	17	28.3	60	6.BD006937	BD006937 Artificia
45	17	28.3	67	6.AR055664	AR055664 Sequence

ALIGNMENTS

RESULT 1	BD006972	Artificial promoter libraries for selected organisms and promoters derived from such libraries.	60 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006972					
DEFINITION	BD006972.1 GI:18635343					
ACCESSION	BD006972.1					
VERSION	UP 2001503249-A/41.					
KEYWORDS	Lactococcus lactis					
SOURCE	Lactococcus lactis					
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.					
REFERENCE	1 (bases 1 to 60)					
AUTHORS	Hammer, K. and Janssen, P. R.					
TITLE	Artificial promoter libraries for selected organisms and promoters derived from such libraries					
JOURNAL	Patent: JP 2001503249-A 41 13-MAR-2001;					
COMMENT	PETER RUDAR JANSSEN.					
OS	Lactococcus lactis					
PN	UP 2001503249-A/41.					
PD	13-MAR-2001					
PF	25-AUG-1997 JP 1998510287					
PR	23-AUG-1996 DK 0886/96					
PI	KALIN HAMMER, PETER RUDAR JANSSEN					
PC	CI2N15/09, CI2N15/00					
CC	Strandedness: Double;					
CC	Topology: linear;					
FT	Key Location/Qualifiers					
FEATURES	promoter (4)..(60).					
source	Location/Qualifiers					
	1..60					
	/organism="Lactococcus lactis"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:1358"					
ORIGIN						
Query Match	100.0%;	Score 60;	DB 6;	Length 60;		
Best Local Similarity	100.0%;	Pred. No. 9.8e-23;				
Matches	60;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
1	CATCTTGTAGTTATCTTGACAAAGTATGAGACTGATATATAGTGTAGTACTGTT 60					

GenCore version 5.1.6
Copyright: (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-40

Perfect score: 60
Sequence: 1 TATGCGTAGTTATCTTG.....TATATGCGTCTAGATTAG 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn190s:*\n3: geneseqn2000s:*\n4: geneseqn2001s:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	2	AAV23212
2	19	31.7	60	2	AAV23188
3	18	30.0	59	2	AAV23209
4	17	28.3	58	2	AAV23184
5	17	28.3	59	2	AAV23193
6	17	28.3	60	2	AAV23197
7	17	28.3	60	6	ABK98622
8	17	28.3	60	9	ACD13873
9	17	28.3	95	6	ABK98612
10	17	28.3	95	6	ACD13863
11	17	28.3	9437	6	ABK98601
12	17	28.3	9437	9	ACD13852
13	17	28.3	10929	6	ABK98591
14	17	28.3	10929	9	ACD13842
15	17	28.3	12739	6	ABK98592
16	17	28.3	12739	9	ACD13843
17	17	28.3	15320	4	ABL07834
18	17	28.3	84428	12	ADM45913
19	17	28.3	130480	4	AAF25833
20	17	28.3	185371	6	ABT10718

21	16	26.7	58	2	AAV23186	AAV23186	Lactococc
22	16	26.7	59	2	AAV23210	AAV23210	Lactococc
23	16	26.7	59	2	AAV23206	AAV23206	Lactococc
24	16	26.7	60	2	AAV23194	AAV23194	Lactococc
25	16	26.7	60	2	AAV23213	AAV23213	Lactococc
26	16	26.7	60	2	AAV23200	AAV23200	Lactococc
27	16	26.7	60	2	AAV23214	AAV23214	Lactococc
28	16	26.7	60	2	AAV23178	AAV23178	Lactococc
29	16	26.7	60	2	AAV23199	AAV23199	Lactococc
30	16	26.7	60	2	AAV23205	AAV23205	Lactococc
31	16	26.7	207	6	ABK98627	ABK98627	L. lactis
32	16	26.7	207	9	ACD13878	ACD13878	L. lactis
33	16	26.7	242	6	ABK98617	ABK98617	L. lactis
34	16	26.7	242	9	ACD13868	ACD13868	L. lactis
35	16	26.7	909	12	ADG70458	ADG70458	Alginic a
36	16	26.7	1881	6	AB214845	AB214845	Arabidops
37	16	26.7	2016	3	AA050686	AA050686	Arabidops
38	16	26.7	9566	6	ABK98634	ABK98634	Vector pB
39	16	26.7	9566	9	ACD13885	ACD13885	L. lactis
40	16	26.7	18061	6	AB161973	AB161973	Colon ade
41	16	26.7	18061	6	AB162480	AB162480	Colon ade
42	16	26.7	18061	6	AB162481	AB162481	Colon ade
43	16	26.7	18061	12	ADQ18416	ADQ18416	Human sof
44	16	26.7	78313	9	ADA02618	ADA02618	Human FYN
45	16	26.7	78313	10	ADB72356	ADB72356	Human FYN

ALIGNMENTS

RESULT 1	AAV23212	standard; DNA; 60 BP.
ID	AAV23212	standard; DNA; 60 BP.
AC	AAV23212;	
XX		
DT	28-JUL-1998	(first entry)
XX		
DE	Lactococcus lactis	constitutional promoter Cp7.
XX		
KW	Lactococcus lactis;	constitutional promoter; optima; spacer;
XX	artificial promoter library;	gene expression; ds.
OS	Synthetic.	
OS	Lactococcus lactis.	
XX		
FT	Key	Location/Qualifiers
FT	promoter	4..60
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
PN	W09807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DX000342.
XX		
PR	23-AUG-1996;	96DX-00000886.
XX		
PA	(JENSEN) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries -	containing consensus promoter
PT	sequences and variable spacers,	used to generate promoters for optimising
PT	expression of genes.	
XX		
PS	Claim 28; Page 59; 89p;	English.
XX		
CC	This is a Lactococcus lactis	constitutional promoter sequence used in the
CC	construction of an artificial	promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 05:53:11 ; Search time 1455.06 Seconds
(without alignments)
1569.593 Million cell updates/sec

Title: US-09-242-657B-41

Perfect score: 60
Sequence: 1 CATCTTTACTTATTTCTG.....TATATAGTGAGTACTGTT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	33.3	854	9	EX137003 Danio rer
2	20	33.3	1934	8	BH770983 LMGtag70
3	19	31.7	478	8	BH339049 CH230-111
4	19	31.7	646	8	AO510836 nbdb0048L
5	19	31.7	661	4	B1597545 603247074
6	18	30.0	358	9	AJ595995 Arabidops
7	18	30.0	374	9	CC793807 SALU_0187
8	18	30.0	417	8	B2156236 CH230-300
9	18	30.0	528	8	CE356637 Cigr-gss-
10	18	30.0	566	2	BF232077 del5a01.x
11	18	30.0	593	8	B2911974 CH240.110
12	18	30.0	602	9	CC956633 B0THJ79TP
13	18	30.0	612	6	CA875995 669609 NC
14	18	30.0	633	7	CK753818 eca01-13c
15	18	30.0	711	8	BH945783 obu91f08.
16	18	30.0	759	8	B2074299 1Kh09d04.
17	18	30.0	809	8	BH676155 BOMB66TF
18	18	30.0	829	8	B2209454 CH230-437
19	17	28.3	275	7	CN775743 ta884h03.
20	17	28.3	333	8	B2958732 PGRH09G19
21	17	28.3	347	7	CN775920 cae79b05.
22	17	28.3	365	8	AO293178 HS_2201_B
23	17	28.3	401	9	CR498859 Medicago
24	17	28.3	422	5	BW204117 BW204117

c	25	17	28.3	454	8	A2122453	RPCI-23-4
c	26	17	28.3	466	7	CO536440	eah7a11.
c	27	17	28.3	466	8	AO724742	HS_5402_A
c	28	17	28.3	524	7	CO338677	ENI7401.5
c	29	17	28.3	551	2	BF111841	7136c07.x
c	30	17	28.3	561	6	CD307372	StcPu691.
c	31	17	28.3	573	8	AO419867	RPCI-11-1
c	32	17	28.3	576	7	CN183605	UCRCS04.0
c	33	17	28.3	597	4	BJ110901	BJ110901
c	34	17	28.3	616	8	AZ558305	RPCI-23-2
c	35	17	28.3	626	7	CF790646	875803_MA
c	36	17	28.3	634	1	AV243888	AV243888
c	37	17	28.3	646	9	CE800258	Cigr-gss-
c	38	17	28.3	650	7	CF363454	831311_MA
c	39	17	28.3	659	7	CF362159	828548_MA
c	40	17	28.3	668	7	CN632467	taf02a04.
c	41	17	28.3	679	9	AG565424	Mus muscu
c	42	17	28.3	681	9	CE805815	Cigr-gss-
c	43	17	28.3	683	7	CN625364	tae01a11.
c	44	17	28.3	685	9	CE127633	Cigr-gss-
c	45	17	28.3	699	5	B0895332	X022506_P

ALIGNMENTS

RESULT 1
EX137003
LOCUS
DEFINITION
Danio rerio genomic clone DKEY-90A17, genomic survey sequence.
ACCESSION
EX137003.1 GI:27968314
VERSION
GSS.
KEYWORDS
EX137003.1
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk unpublished
Direct Submission
1 (bases 1 to 854)
Humphray, S.J., Huckle, E. and Durham, J.L.
Cytpriniformes; Cyprinidae; Danio.

COMMENT
This sequence was generated from the SP6 end of BAC 90A17. 90A17 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Kysene. Further details:
http://www.sanger.ac.uk/Projects/D_rerio/
Location/Qualifiers

FEATURES

Source
1..854
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-90A17"
/issue_type="Testis"
/note="vector pindigoBAC-536"

ORIGIN

Query Match 33.3% Score 20; DB 9; Length 854;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTTAGTTTATCTTGACAA 25
|||||
Db 368 TTTAGTTTATCTTGACAA 387

RESULT 2
BH770983/c
LOCUS
DEFINITION
LMGtag708 MG3163 Random Sequence Tag Library Lactococcus lactis subsp. cremoris genomic, genomic survey sequence.
ACCESSION
BH770983

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 293.652 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-42

Perfect score: 60
Sequence: 1 CATAGCTAGTTATCTCTG.....TATAATATTAGTACTGT 60

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:.*
1: gb ha:.*
2: gb hgt:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_scs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	BD006973	BD006973 Artificial
2	60	33.3	187665	10 AL772301	AL772301 Mouse DNA
3	20	33.3	209864	2 AC121779	AC121779 Mus muscu
4	19	31.7	67	6 AR055663	AR055663 Sequence
5	19	31.7	1100	1 LTRRRM	XR7564 L.lactis ge
6	19	31.7	1100	6 AR055662	AR055662 Sequence
7	19	31.7	12737	1 AE006241	AE006241 Lactococc
8	19	31.7	143396	5 BX294186	BX294186 Zebrafish
9	19	31.7	146570	3 AC117076	AC117076 Dictyoste
10	19	31.7	151859	9 AC023405	AC023405 Homo sapi
11	19	31.7	175745	9 AC103784	AC103784 Homo sapi
12	18	30.0	4140	6 AR551272	AR551272 Sequence
13	18	30.0	5693	6 CQ602957	CQ602957 Sequence
14	18	30.0	10639	2 AC017273	AC017273 Drosophi
15	18	30.0	49321	10 AF275366	AF275366 Mus muscu
16	18	30.0	91685	2 AC005121	AC005121 Drosophi
17	18	30.0	112702	2 DMR2962	DMR2962 Drosophi
18	18	30.0	129595	8 AP005808	AP005808 Oryza sat
19	18	30.0	142915	9 AC105288	AC105288 Homo sapi

20	18	30.0	144983	2 AP004793	AP004793 Oryza sat
21	18	30.0	145120	8 AP003434	AP003434 Oryza sat
22	18	30.0	161348	10 AL645532	AL645532 Mouse DNA
23	18	30.0	163620	10 AC133892	AC133892 Mus muscu
24	18	30.0	170362	3 AC009250	AC009250 Drosophi
25	18	30.0	177000	2 AC139021	AC139021 Mus muscu
26	18	30.0	181413	5 C0870098	C0870098 Sequence
27	18	30.0	201021	5 AL954702	AL954702 Zebrafish
28	18	30.0	209859	5 BX005398	BX005398 Zebrafish
29	18	30.0	211716	2 AC124786	AC124786 Mus muscu
30	18	30.0	243819	2 AC107341	AC107341 Rattus no
31	18	30.0	257960	2 AC129833	AC129833 Rattus no
32	18	30.0	262525	3 AB003652	AB003652 Drosophi
33	18	30.0	276559	2 AC095403	AC095403 Rattus no
34	17	28.3	60	6 BD006947	BD006947 Artificial
35	17	28.3	60	6 BD006956	BD006956 Artificial
36	17	28.3	60	6 BD006964	BD006964 Artificial
37	17	28.3	1332	8 SCRP15B	X51520 Yeast Yrp1
38	17	28.3	2000	6 AX654949	AX654949 Sequence
39	17	28.3	3726	9 AK095111	AK095111 Homo sapi
40	17	28.3	3726	9 AK095111	AK095111 Homo sapi
41	17	28.3	6160	6 AX344269	AX344269 Sequence
42	17	28.3	6160	6 AX348666	AX348666 Sequence
43	17	28.3	6483	3 D16579	D16579 Dictyoste
44	17	28.3	9863	6 C0600818	C0600818 Sequence
45	17	28.3	10774	9 AB067481	AB067481 Homo sapi

ALIGNMENTS

RESULT 1
BD006973
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION
BD006973
VERSION
JP 2001503249-A/42.
KEYWORDS
Lactococcus lactis
SOURCE
Lactococcus lactis
ORGANISM
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

REFERENCE
1 (bases 1 to 60)

AUTHORS
Hammer, K. and Janssen, P. R.

TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries

JOURNAL
Patent: JP 2001503249-A 42 13-MAR-2001;

COMMENT
PETER RUDAR JANSSEN

OS Lactococcus lactis
PN JP 2001503249-A/42
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter
Location/Qualifiers
(4) . (60)
1..60
/organism="Lactococcus lactis"
/mol_type="genomic DNA"
/db_xref="taxon:1358"

ORIGIN

Query Match 100.0%; Score: 60; DB 6; Length 60;
Best local similarity 100.0%; Pred. No. 1.2e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATAGCTAGTTATCTTGAACGCGTCATGCTGGTATATATTAGTACTGT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2005, 18:42:28 ; Search time 198.929 Seconds
(without alignments)
1785.485 Million cell updates/sec

Title: US-09-242-657B-41

Perfect score: 60
Sequence: 1 CATCTTACTTACTCTG.....TATTAAGCTGACTCTT 60

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	60	100.0	60 2 AAV23213	AAV23213 Lactococc
2	23	38.3	60 2 AAV23200	AAV23200 Lactococc
3	18	30.0	60 2 AAV23205	AAV23205 Lactococc
4	18	30.0	207 6 ABK98627	ABK98627 L. lactis
5	18	30.0	207 6 ACD13878	ACD13878 L. lactis
6	18	30.0	242 6 ABK98617	ABK98617 L. lactis
7	18	30.0	242 6 ACD13868	ACD13868 L. lactis
8	18	30.0	9566 6 ABK98634	ABK98634 Vector pr
9	18	30.0	9566 6 ACD13885	ACD13885 L. lactis
10	18	30.0	110000 6 ABA90521.00	ABA90521 Genomic s
11	17	28.3	59 2 AAV23209	AAV23209 Lactococc
12	17	28.3	60 2 AAV23178	AAV23178 Lactococc
13	17	28.3	115 6 ABK98626	ABK98626 L. lactis
14	17	28.3	115 6 ACD13877	ACD13877 L. lactis
15	17	28.3	150 6 ABK98616	ABK98616 L. lactis
16	17	28.3	150 6 ACD13867	ACD13867 L. lactis
17	17	28.3	9484 6 ABK98633	ABK98633 Vector pl
18	17	28.3	9484 6 ACD13884	ACD13884 L. lactis
19	17	28.3	12216 4 AAS334427	AAS334427 DNA encod
20	17	28.3	46404 11 ACN44270	ACN44270 Human gen

21	17	28.3	107829	11	ACN44088	ACN44088 Mouse gen
22	16	26.7	58	2	AAV23184	AAV23184 Lactococc
23	16	26.7	60	2	AAV23188	AAV23188 Lactococc
24	16	26.7	60	2	AAV23197	AAV23197 Lactococc
25	16	26.7	60	2	AAV23214	AAV23214 Lactococc
26	16	26.7	60	2	AAV23212	AAV23212 Lactococc
27	16	26.7	265	2	AAQ23880	AAQ23880 pKTH1820
28	16	26.7	265	2	AAQ46205	AAQ46205 Promoter
29	16	26.7	265	2	AAT31875	AAT31875 Promoter
30	16	26.7	460	9	ACH27886	ACH27886 Human adu
31	16	26.7	461	8	ABX40614	ABX40614 Bovine ES
32	16	26.7	835	8	ABZ35933	ABZ35933 Human sec
33	16	26.7	841	4	AAQ07905	AAQ07905 Human sec
34	16	26.7	1839	4	ABJ29475	ABJ29475 Drosophil
35	16	26.7	1944	5	ABV24919	ABV24919 Human pro
36	16	26.7	3120	13	ADR06772	ADR06772 Full leng
37	16	26.7	3930	4	AAI64767	AAI64767 Human cto
38	16	26.7	4241	4	ABJ29474	ABJ29474 Drosophil
39	16	26.7	4648	13	ADR06982	ADR06982 Full leng
40	16	26.7	14301	4	ABJ02084	ABJ02084 Drosophil
41	16	26.7	14770	5	AAF24298	AAF24298 Legionell
42	16	26.7	14770	6	ABK10464	ABK10464 Legionell
43	16	26.7	88892	12	ADQ97695	ADQ97695 Human can
44	16	26.7	110000	6	ABA90521.02	ABA90521.02 Continuation (3 of
45	16	26.7	117754	11	ACN43866	ACN43866 Human gen

ALIGNMENTS

RESULT 1	AAV23213	standard, DNA; 60 BP.
ID	AAV23213	
AC	AAV23213;	
XX		
DT	28-JUL-1998	(first entry)
XX		
DE	Lactococcus lactis constitutional promoter Cp8.	
XX		
KW	Lactococcus lactis; constitutional promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; ds.	
OS	Synthetic.	
OS	Lactococcus lactis.	
XX		
FH	Key	Location/Qualifiers
FT	Promoter	4..60
FT		/*tag= a
FT		/standard_name= "Constitutional promoter"
XX		
PN	WO9807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PP	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENSEN) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
PS	Claim 28; Page 60; 89pp; English.	
XX		
CC	This is a Lactococcus lactis constitutional promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 494.595 Seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-42

Perfect score: 60
Sequence: 1 CATAGCTAGTATTACTCTG.....TATTAATTAATTAAGTACTGT 60

Scoring table: OLIGO_NUC

Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	31.7	256	2	BB053672 BB053672
2	19	31.7	658	1	AI667210 f014c07.x
3	18	30.0	293	2	BB337900 BB337900
4	18	30.0	310	9	CE195770 CE195770
5	18	30.0	338	5	BM079141 BM079141
6	18	30.0	360	5	BM214879 BM214879
7	18	30.0	449	5	EX561394 EX561394
8	18	30.0	591	4	BG524755 BG524755
9	18	30.0	602	4	BG523715 34-73 Stev
10	18	30.0	606	4	BG524506 43-32 Ste
11	18	30.0	642	4	BG522350 20-48 Ste
12	18	30.0	661	7	CO638156 USDA-FP 1
13	18	30.0	674	4	BG522767 25-30 Ste
14	18	30.0	705	9	CR504842 Medicago
15	18	30.0	714	6	CR436148 616469 MA
16	18	30.0	743	9	CR502660 Medicago
17	18	30.0	890	9	CG952586 MBE17937F
18	17	28.3	340	8	CC170014 1194f10.9
19	17	28.3	368	1	CE159974 figt-gss-
20	17	28.3	382	1	AI329814 Bf04ane.r
21	17	28.3	385	6	CD197114 MSI-0094P
22	17	28.3	396	2	CD421133 AUF-IPSPN
23	17	28.3	399	2	BF661850 maa7ab10.
24	17	28.3	414	2	BF661229 maa7ab10.

25	17	28.3	432	5	BU014513
26	17	28.3	446	5	CU775176
27	17	28.3	455	5	BY424520
28	17	28.3	460	2	AW835252
29	17	28.3	527	5	BO559914
30	17	28.3	542	4	BI616968
31	17	28.3	554	1	AU722507
32	17	28.3	574	4	BM215328
33	17	28.3	618	4	BI677276
34	17	28.3	618	4	BI615973
35	17	28.3	626	9	CM521005
36	17	28.3	648	4	BM540077
37	17	28.3	667	7	CO201140
38	17	28.3	668	9	CU551243
39	17	28.3	683	9	CE683173
40	17	28.3	685	1	AV256717
41	17	28.3	697	8	BH929272
42	17	28.3	726	8	CC170012
43	17	28.3	786	9	CU580107
44	17	28.3	806	2	BP064511
45	17	28.3	816	5	BU747545

ALIGNMENTS

RESULT 1
BB053672 256 bp mRNA linear EST 25-JUN-2000
LOCUS BB053672.1 RIKEN full-length enriched, 12 days embryo male wolffian
DEFINITION BB053672.1 RIKEN full-length enriched, 12 days embryo male wolffian
KEYWORDS BB053672.1 GI:8460820
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB053672 256 bp mRNA linear EST 25-JUN-2000
BB053672.1 RIKEN full-length enriched, 12 days embryo male wolffian
BB053672.1 GI:8460820
EST.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 256)
Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C.,
Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H.,
Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suganara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,
Takahashi, F., Tomimaga, N., Toyota, T., Tsunoda, Y., Watanabe, A.,
Watanabe, S., Yamamura, T., Yamashita, T., Yano, R., Yasunishi, A.,
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
RIKEN Mouse ESTs (Komno, H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermotabilization and thermoactivation of thermostable enzymes by
trichostatin and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunari, T., Akiyama, U., Shibata, K., Izawa, M., Kawai, J.,
Tomari, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 / Search time 866.273 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-43

Sequence: 1 GAATTCGACTCAACGCG.....TCGCTACCATCATGATCC 177

Scoring table: OLIGO_NUC

Gapop 60.0, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues.

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenBml:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	177	100.0	177	6	BD006974
2	177	43.5	195	6	BD006986
3	76	42.9	167	6	BD006977
4	76	42.9	195	6	BD006979
5	76	42.9	195	6	BD006984
6	76	42.9	199	6	BD006934
7	71	40.1	184	8	SCYOL140W
8	71	40.1	9892	8	SCARGCDC
9	68	38.4	189	6	BD006987
10	68	38.4	195	6	BD006981
11	68	38.4	2479	8	SCYOL141W
12	65	36.7	193	6	BD006982
13	58	32.8	176	6	BD006980
14	51	28.8	179	6	BD006989
15	51	28.8	195	6	BD006988
16	48	27.1	191	6	BD006978
17	45	26.0	191	6	BD006976
18	43	24.3	188	6	BD006985
19	40	22.6	166	6	BD006983

20	25	14.1	1002	11	CNS06JTO	AL402010 T7 end of
21	23	13.0	182	6	BD006975	BD006975 Artificia
22	21	11.9	127376	2	AC141672	AC141672 Apis mell
23	21	11.9	196217	2	AC131340	AC131340 Mus muscu
24	20	11.3	143255	2	AC015996	AC015996 Homo sapi
25	20	11.3	153887	9	AC093830	AC093830 Homo sapi
26	20	11.3	199465	9	AC087664	AC087664 Homo sapi
27	20	11.3	237468	2	AC108532	AC108532 Rattus no
28	20	11.3	309838	2	AC118433	AC118433 Rattus no
29	19	10.7	1876	3	AF300529	AF300529 Diadadia
30	19	10.7	11000	1	AE017308-4	Continuation (5 of
31	19	10.7	129675	2	AC102950	AC102950 Homo sapi
32	19	10.7	179503	9	AL445468	AL445468 Human DNA
33	19	10.7	199814	3	CEX47D3A	AL117202 Caenorhab
34	19	10.7	204800	2	AC120282	AC120282 Rattus no
35	19	10.7	214074	2	AC148841	AC148841 Gorilla g
36	19	10.7	214455	2	AC118451	AC118451 Rattus no
37	19	10.7	216113	2	AC116121	AC116121 Mus muscu
38	18	10.2	2498	1	BTHKNA	U03552 Bacillus th
39	18	10.2	5082	10	RNCFTFR3UT	AF224431 Rattus no
40	18	10.2	5198	6	AX345180	AX345180 Sequence
41	18	10.2	6059	6	AX281301	AX281301 Sequence
42	18	10.2	11051	1	AE013911	AE013911 Yersinia
43	18	10.2	11951	1	AE010615	AE010615 Pseudocete
44	18	10.2	14011	1	AE001329	AE001329 Chlamydia
45	18	10.2	14720	5	AB092690	AB092690 Oryzias 1

ALIGNMENTS

RESULT 1
BD006974 177 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION BD006974.1 GI:18635345
VERSION JP 2001503249-A/43
KEYWORDS
SOURCE JP 2001503249-A/43
ORGANISM
Saccharomyces cerevisiae (baker's yeast)
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
REFERENCE
1 (bases 1 to 177)
Hämmer, K. and Janssen, P.R.
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
Patent: JP 2001503249-A 43 13-MAR-2001;
PETER RUDAR JANSSEN
COMMENT
OS Saccharomyces cerevisiae (yeast)
PN JP 2001503249-A/43
PD 13-MAR-2001 JP 1998510287
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter
Location/Qualifiers
(8). (177).
1. 177
Location/Qualifiers
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 177; DB 6; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.3e-85;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGACTCAACGCGGTGCTGACGCGGTGCTCAATTAAATGCGCTCTTAT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 54.6803 Seconds

(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-42

Sequence: 1 CATAGCTGATTATCTTG.....TATTAATATTAGTACTGT 60

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneeqn1980s:*\n2: geneeqn1980s:*\n3: geneeqn2000s:*\n4: geneeqn2001as:*\n5: geneeqn2001bs:*\n6: geneeqn2002as:*\n7: geneeqn2002bs:*\n8: geneeqn2003as:*\n9: geneeqn2003bs:*\n10: geneeqn2003cs:*\n11: geneeqn2003ds:*\n12: geneeqn2004as:*\n13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	60	AAV23214	AAV23214 Lactococc
2	19	31.7	207	ABK98627	Abk98627 L. lactis
3	19	31.7	207	ACD13878	Acid13878 L. lactis
4	19	31.7	242	ABK98617	Abk98617 L. lactis
5	19	31.7	242	ACD13868	Acid13868 L. lactis
6	19	31.7	9566	ABK98634	Abk98634 Vector PB
7	19	31.7	9566	ACD13885	Acid13885 L. lactis
8	19	31.7	110000	6 ABA90521_00	Abk90521 Genomic B
9	18	30.0	5693	4 AB122316	Ab122316 Drosophila
10	18	30.0	181413	13 ABD32851	Abd32851 Mouse can
11	17	28.3	60	2 AAV23188	AAV23188 Lactococc
12	17	28.3	60	2 AAV23197	AAV23197 Lactococc
13	17	28.3	60	2 AAV23205	AAV23205 Lactococc
14	17	28.3	310	4 AAK56355	AAK56355 Human imm.
15	17	28.3	361	4 AAK57872	AAK57872 Human imm.
16	17	28.3	636	13 ADO51279	Ado51279 Novel can
17	17	28.3	3726	8 ADA71496	Ada71496 Rice gene
18	17	28.3	3726	8 ADM01848	Adm01848 Human CDN
19	17	28.3	3830	12 ADH22415	Adh22415 Human CDN
20	17	28.3	6160	6 ABK31273	Abk31273 Signal tr

ALIGNMENTS

21	17	28.3	6160	6	ABL70234	Ab170234 Chemical
22	17	28.3	9863	4	AB120890	Ab120890 Drosophila
23	17	28.3	10203	4	AAK90837	AAK90837 Human dig
24	17	28.3	13673	4	AB116734	Ab116734 Drosophila
25	17	28.3	12767	13	ABD32657	Abd32657 Mouse can
26	16	26.7	58	2	AAV23184	AAV23184 Lactococc
27	16	26.7	59	2	AAV23190	AAV23190 Lactococc
28	16	26.7	59	2	AAV23199	AAV23199 Lactococc
29	16	26.7	60	2	AAV23211	AAV23211 Lactococc
30	16	26.7	60	2	AAV23182	AAV23182 Lactococc
31	16	26.7	60	2	AAV23208	AAV23208 Lactococc
32	16	26.7	60	2	AAV23213	AAV23213 Lactococc
33	16	26.7	60	2	AAV23212	AAV23212 Lactococc
34	16	26.7	60	2	AAV23178	AAV23178 Lactococc
35	16	26.7	60	2	AAV23180	AAV23180 Lactococc
36	16	26.7	419	12	ADP95471	Adp95471 Cotton ex
37	16	26.7	597	10	ADC26024	Adc26024 Maize met
38	16	26.7	597	12	ADJ94393	Adj94393 Maize met
39	16	26.7	609	8	AA160099	AA160099 Maize met
40	16	26.7	612	3	AAK97281	AAK97281 CDNA enco
41	16	26.7	651	3	AA113601	AA113601 Aspergill
42	16	26.7	834	6	ACA47045	ACA47045 Prokaryot
43	16	26.7	840	8	ABN91922	Abn91922 Staphyloc
44	16	26.7	840	13	AD500782	Ad500782 Staphyloc
45	16	26.7	1247	13	ADT05082	Adt05082 Haemophil

RESULT 1
AAV23214
ID AAV23214 standard; DNA, 60 BP.

AAV23214;

28-JUL-1998 (first entry)

Lactococcus lactis constitutional promoter Cp9.

Lactococcus lactis; constitutional promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Synthetic.

Lactococcus lactis.

Key Location/Qualifiers

promoter 4..60

W09807846-A1.

26-FEB-1998.

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENSEN) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 60; 89pp; English.

This is a Lactococcus lactis constitutional promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1459.05 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-43

Perfect score: 177

Sequence: 1 GAATTCGTCGACCTCAACGGG.....TCGCTACCATCATGATCC 177

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	11.3	450	2	AM003938
C 2	20	11.3	775	2	CL582491
C 3	20	11.3	1145	6	CB565228
C 4	19	10.7	601	7	CK556464
C 5	19	10.7	742	8	BH966805
C 6	19	10.7	1220	9	AG035297
C 7	18	10.2	228	6	CB449507
C 8	18	10.2	235	6	CB440185
C 9	18	10.2	235	6	CB451827
C 10	18	10.2	397	8	AZ146750
C 11	18	10.2	407	8	CC153956
C 12	18	10.2	443	7	CR471577
C 13	18	10.2	451	8	CC094891
C 14	18	10.2	452	9	CC0842207
C 15	18	10.2	469	8	BZ950076
C 16	18	10.2	502	8	BH004907
C 17	18	10.2	578	5	BP102832
C 18	18	10.2	593	8	AZ255847
C 19	18	10.2	654	5	BM107202
C 20	18	10.2	658	5	BY727726
C 21	18	10.2	660	5	BM068182
C 22	18	10.2	664	5	BM121165
C 23	18	10.2	681	9	CL750452
C 24	18	10.2	683	5	BM107288

25	18	10.2	716	9	BM184773	BM184773	Danio rer
C 26	18	10.2	716	9	CC860315	CC860315	NDL.11351
C 27	18	10.2	723	9	BM187483	BM187483	Danio rer
28	18	10.2	755	9	BM160384	BM160384	Danio rer
29	18	10.2	760	9	CC874751	CC874751	NDL.36A22
30	18	10.2	763	9	BM131577	BM131577	Danio rer
C 31	18	10.2	778	9	CL843043	CL843043	OR_CBA007
C 32	18	10.2	795	9	CC870420	CC870420	NDL.12908
C 33	18	10.2	801	8	CC150321	CC150321	NDL.60C14
C 34	18	10.2	811	9	CC867413	CC867413	NDL.1715
C 35	18	10.2	831	9	CC500494	CC500494	CH240.338
36	18	10.2	832	7	CK603478	CK603478	AGENCOURT
37	18	10.2	869	9	CL684606	CL684606	FR1013A_B
C 38	18	10.2	879	9	CC851624	CC851624	NDL.124E2
C 39	18	10.2	893	9	CC847477	CC847477	NDL.22G12
C 40	18	10.2	913	8	CC068647	CC068647	CSU-K33r
41	18	10.2	958	8	CC109175	CC109175	NDL.51H12
42	18	10.2	1008	9	CNS0370X	AL231666	Tetradon
C 43	18	10.2	1142	8	BZ558780	BZ558780	pa98401_4
C 44	18	10.2	1162	8	CC294522	CC294522	CH261-43F
C 45	18	10.2	1433	8	BZ571137	BZ571137	msb2_1745

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
AM003938
450 bp mRNA linear EST 08-SRP-1999
wg84f04.x1 NCI_CGAP_GC6 Homo sapiens CDNA clone IMAGE:2478751 3'

AM003938
Accession
Version
Keywords
Source
Organism
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
NCI-CGAP
http://www.ncbi.nlm.nih.gov/ncicgap.
1 (bases 1 to 450)

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNU at:
www-bio.lnl.gov/bbcp/image/image.html
Seq primer: -400P from Gibco.

FEATURES

source
Location/Qualifiers
1..450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pRTD-Pac (Pharmacia) with a modified polylinker. Site 1: Not I, Site 2: Eco RI, Plasmid DNA from the normalized library NCI-CGAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470993, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 161.307 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-43

Perfect score: 177
Sequence: 1 GAATTCGTGACTCAACGGG.....TCGTCACCAATGATGATCC 177

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	177	100.0	177	2	AAV23215 Saccharom
2	77	43.5	195	2	AAV23227 Saccharom
3	76	42.9	167	2	AAV23218 Saccharom
4	76	42.9	195	2	AAV23225 Saccharom
5	76	42.9	195	2	AAV23220 Saccharom
6	76	42.9	199	2	AAV23176 Saccharom
7	68	38.4	189	2	AAV23228 Saccharom
8	68	38.4	195	2	AAV23222 Saccharom
9	65	36.7	193	2	AAV23223 Saccharom
10	58	32.8	179	2	AAV23221 Saccharom
11	51	28.8	176	2	AAV23230 Saccharom
12	51	28.8	195	2	AAV23229 Saccharom
13	48	27.1	191	2	AAV23219 Saccharom
14	46	26.0	191	2	AAV23217 Saccharom
15	40	22.6	166	2	AAV23224 Saccharom
16	39	22.0	188	2	AAV23226 Saccharom
17	23	13.0	182	2	AAV23216 Saccharom
18	18	10.2	5198	6	ABL32278 Human Imm
19	18	10.2	6059	6	ABL54343 Chemical1
20	18	10.2	110000	2	AAZ01425_08 Continuation (9 of

ALIGNMENTS

21	17	9.6	43	12	ADH42842	Adh42842 Novel hum
22	17	9.6	498	12	ADL1516	Adl1516 Cat flea
23	17	9.6	576	6	ABO44458	Abg44458 Oligonuc
24	17	9.6	576	6	ABO44459	Abg44459 Oligonuc
25	17	9.6	1866	13	AD847979	Ad847979 Bacterial
26	17	9.6	1959	6	AA27607	AA27607 Synchocy
27	17	9.6	2000	6	AB215437	Ab215437 Arabidops
28	17	9.6	2000	6	ADA68824	Ada68824 Arabidops
29	17	9.6	2789	6	ABO70842	Abg70842 listeria
30	17	9.6	5326	4	ABL10170	Ab110170 Drosophi
31	17	9.6	5572	6	ABL32815	Ab132815 Human imm
32	17	9.6	6000	6	ABL33133	Ab133133 Human imm
33	17	9.6	6160	6	ABK31273	Abk31273 Signal tr
34	17	9.6	6160	6	ABL70234	Ab170234 Chemical1
35	17	9.6	6163	6	ABN80118	Abn80118 Human che
36	17	9.6	6219	6	ABL32866	Ab132866 Human imm
37	17	9.6	6219	6	AA63324	AA63324 Human imm
38	17	9.6	6447	6	ABL33086	Ab133086 Human met
39	17	9.6	6447	6	ABL34532	Ab134532 Human met
40	17	9.6	6447	6	ABL70547	Ab170547 Chemical1
41	17	9.6	6447	6	ABO65973	Abg65973 Human ang
42	17	9.6	6447	7	ADS9793	Ads9793 Bistulphic
43	17	9.6	8305	6	ABL33568	Ab133568 Human imm
44	17	9.6	10872	6	ABK31389	Abk31389 Signal tr
45	17	9.6	10872	6	ABL70336	Ab170336 Chemical1

RESULT 1

ID AAV23215 standard; DNA, 177 BP.

AAV23215;

28-JUL-1998 (first entry)

Saccharomyces cerevisiae promoter Yp102.

Saccharomyces cerevisiae; yeast promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Saccharomyces cerevisiae.

OS Saccharomyces cerevisiae.

Key Location/Qualifiers

FT Promoter 8..177

FT Promoter /*tag=a

FT Promoter /standard_name="Yeast promoter"

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENS) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 61; 89pp; English.

This is a Saccharomyces cerevisiae promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 165.864 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-44

Perfect score: 182

Sequence: 1 GAATTCGTACTCAGCGCAT.....TCGCTACCAATCATGATCC 182

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 segs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	100.0	182	2	AAV23216
2	182	100.0	182	2	AAV23216
3	182	100.0	182	2	AAV23216
4	182	100.0	182	2	AAV23216
5	182	100.0	182	2	AAV23216
6	182	100.0	182	2	AAV23216
7	182	100.0	182	2	AAV23216
8	182	100.0	182	2	AAV23216
9	182	100.0	182	2	AAV23216
10	182	100.0	182	2	AAV23216
11	182	100.0	182	2	AAV23216
12	182	100.0	182	2	AAV23216
13	182	100.0	182	2	AAV23216
14	182	100.0	182	2	AAV23216
15	182	100.0	182	2	AAV23216
16	182	100.0	182	2	AAV23216
17	182	100.0	182	2	AAV23216
18	182	100.0	182	2	AAV23216
19	182	100.0	182	2	AAV23216
20	182	100.0	182	2	AAV23216

21	18	9.9	455	9	ACH36676
22	18	9.9	550	6	ABQ15036
23	18	9.9	550	6	ABQ15037
24	18	9.9	641	4	AAS27518
25	18	9.9	641	4	ABK43839
26	18	9.9	641	10	ADB93696
27	18	9.9	641	12	AD154226
28	18	9.9	647	13	AD154226
29	18	9.9	924	10	AD153916
30	18	9.9	939	12	AD153916
31	18	9.9	1189	3	AA930372
32	18	9.9	1217	13	ACN41461
33	18	9.9	1219	2	AAV44326
34	18	9.9	1316	13	ACN43085
35	18	9.9	1346	3	AA240664
36	18	9.9	1424	13	ACN41460
37	18	9.9	1443	13	ACN43084
38	18	9.9	1449	3	AA240663
39	18	9.9	1569	6	AA562680
40	18	9.9	1597	6	ABK52766
41	18	9.9	1597	6	ABQ91931
42	18	9.9	1597	10	ADD89980
43	18	9.9	1600	4	ABK43529
44	18	9.9	1600	12	AD153916
45	18	9.9	1696	4	ABK43740

ALIGNMENTS

RESULT 1	AAV23216	standard; DNA; 182 BP.
ID	AAV23216	standard; DNA; 182 BP.
AC	AAV23216	
XX	AAV23216	
DT	28-JUL-1998	(first entry)
XX		
DE	Saccharomyces cerevisiae promoter Yp112.	
XX		
KW	Saccharomyces cerevisiae; yeast promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; ds.	
OS	Synthetic.	
XX	Saccharomyces cerevisiae.	
XX		
FT	Key	Location/Qualifiers
FT	Promoter	8..181
FT		/*tag= a
FT		/standard_name= "Yeast promoter"
XX		
PN	WO9807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENSEN) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
PS	Claim 28, Page 61, 89pp; English.	
XX		
CC	This is a Saccharomyces cerevisiae promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 890.744 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657b-44

Sequence: 1 GAATTCGATCAGCGGATC.....TCGCTACCATCATGATCC 182

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenBank: 1: gb_ha: 2: gb_hg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sta: 12: gb_sy: 13: gb_un: 14: gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182	100.0	182	6	BD006975 Att1ficia
2	23	12.6	166	6	BD006983 Att1ficia
3	23	12.6	167	6	BD006977 Att1ficia
4	23	12.6	176	6	BD006989 Att1ficia
5	23	12.6	177	6	BD006974 Att1ficia
6	23	12.6	179	6	BD006980 Att1ficia
7	23	12.6	188	6	BD006985 Att1ficia
8	23	12.6	189	6	BD006987 Att1ficia
9	23	12.6	191	6	BD006976 Att1ficia
10	23	12.6	191	6	BD006978 Att1ficia
11	23	12.6	193	6	BD006982 Att1ficia
12	23	12.6	195	6	BD006979 Att1ficia
13	23	12.6	195	6	BD006981 Att1ficia
14	23	12.6	195	6	BD006984 Att1ficia
15	23	12.6	195	6	BD006986 Att1ficia
16	23	12.6	195	6	BD006988 Att1ficia
17	23	12.6	199	6	BD006934 Att1ficia
18	21	11.5	1002	11	CNS06JTO AL402010 T7 end of
19	21	11.5	1844	8	SCYOL140W Z74882 S.cerevisia

20	21	11.5	2479	8	SCYOL141W	Z74883 S.cerevisia
21	21	11.5	9892	8	SCARGDC	X84036 S.cerevisia
22	20	11.0	11051	1	AB013911	AB013911 Yersinia
23	20	11.0	16004	3	AY687866	AY687866 Pteronarc
24	20	11.0	99322	2	AC046184	AC046184 Homo sapi
25	20	11.0	103188	8	AP003907	AP003907 Oryza sat
26	20	11.0	110000	1	BX936398_35	Continuation (36 o
27	20	11.0	149995	9	AC016722	AC016722 Homo sapi
28	20	11.0	150614	2	AC074062	AC074062 Homo sapi
29	20	11.0	165416	9	AC034103	AC034103 Homo sapi
30	20	11.0	174001	2	BX927081	BX927081 Danto rer
31	20	11.0	198700	5	AL807378	AL807378 zebrafish
32	20	11.0	199145	2	CR759742	CR759742 Danto rer
33	20	11.0	210050	1	BJ414145	BJ414145 Yersinia
34	20	11.0	291817	1	AB017137	AB017137 Yersinia
35	19	10.4	10592	1	AB013686	AB013686 Yersinia
36	19	10.4	40572	3	U39645	U39645 Caenorhabd
37	19	10.4	95739	2	AP004644	AP004644 Oryza sat
38	19	10.4	96877	2	AC017776	AC017776 Drosophill
39	19	10.4	101799	8	AC099401	AC099401 Oryza sat
40	19	10.4	110000	1	BX936398_09	Continuation (10 o
41	19	10.4	129675	2	AC102950	AC102950 Homo sapi
42	19	10.4	144746	2	AC116867	AC116867 Mus muscu
43	19	10.4	159972	2	AC121554	AC121554 Mus muscu
44	19	10.4	162345	3	AC010659	AC010659 Drosophill
45	19	10.4	165608	2	AC087842	AC087842 Rattus no

ALIGNMENTS

RESULT 1
BD006975
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006975
BD006975 182 bp DNA linear PAT 31-JAN-2002

ACCESSION
BD006975.1 GI:18635346
VERSION
JP 2001503249-A/44.
KEYWORDS
Saccharomyces cerevisiae (baker's yeast)

SOURCE
Saccharomyces cerevisiae
ORGANISM
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE
1. (bases 1 to 182)
Hammer,K. and Janssen,P.R.

AUTHORS
TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 44 13-MAR-2001;

COMMENT
PETER RUDAR JANSSEN
OS
Saccharomyces cerevisiae (yeast)

PN
JP 2001503249-A/44
PD
13-MAR-2001
PR
25-AUG-1997 JP 1998510287

PI
KALIN HAMMER, PETER RUDAR JANSSEN
PC
CI2N15/09, CI2N15/00
CC
Strandedness: Double;

CC
Topology: linear;
FH
Key
FT
promoter (8). (181).

FEATURES
Location/Qualifiers
1..182
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN
source

Query Match

Best Local Similarity 100.0%; Score 182; DB 6; Length 182;
Best Local Similarity 100.0%; Pred. No. 1.5e-90;

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGATCAGCGGATCAGTGGTGGACATGTCAGAACATGTCGAGTTC 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1500.27 Seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-44

Perfect score: 182

Sequence: 1 GAATTCGTACATCATCGCAT.....TCGCTACCATCATGATCC 182

Scoring table: OLIGO_NTC

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database:

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	11.5	650	7	CO793602 NT018B.B1
2	20	11.0	450	2	AM003938
3	20	11.0	687	2	BE287751
4	20	11.0	767	8	AQ747888
5	19	10.4	327	7	CR290121
6	19	10.4	419	9	CE090364
7	19	10.4	449	5	BO970086
8	19	10.4	449	9	CE305098
9	19	10.4	585	9	CE305098
10	19	10.4	694	9	CR251887
11	19	10.4	781	6	CB621855
12	19	10.4	930	8	CC006026
13	19	10.4	1635	9	CL960998
14	18	9.9	103	1	AA002538
15	18	9.9	111	7	TI1445
16	18	9.9	111	7	TI1445
17	18	9.9	114	6	CB866280
18	18	9.9	180	1	AA782682
19	18	9.9	188	1	AA889137
20	18	9.9	193	4	BG955275
21	18	9.9	197	2	AM130521
22	18	9.9	234	6	C75645
23	18	9.9	259	1	AI277671
24	18	9.9	259	9	CG572535

25	18	9.9	262	1	AA889144	AA889144 ak25b09.s
26	18	9.9	265	6	C75462	C75462 C75462 Huma
27	18	9.9	268	6	C75421	C75421 C75421 Huma
28	18	9.9	272	2	BE374505	BE374505 MRI-ST011
29	18	9.9	272	6	C75384	C75384 C75384 Huma
30	18	9.9	276	1	AA781048	AA781048 aj10f04.s
31	18	9.9	282	1	AA579990	AA579990 n179a08.s
32	18	9.9	289	1	AA514325	AA514325 n56b05.s
33	18	9.9	292	2	AM083086	AM083086 xb72g10.x
34	18	9.9	295	1	AA598462	AA598462 ae48b02.s
35	18	9.9	299	5	BU619551	BU619551 UI-H-FH1-
36	18	9.9	303	2	BE467689	BE467689 hz67h09.x
37	18	9.9	303	5	BU621104	BU621104 UI-H-FH1-
38	18	9.9	304	2	BE670821	BE670821 Te4e12.x
39	18	9.9	312	2	AM070858	AM070858 xa05b08.x
40	18	9.9	314	9	CG69367	CG69367 OST46613
41	18	9.9	318	2	BF055500	BF055500 7f73e12.x
42	18	9.9	323	2	BE857977	BE857977 7f73e12.x
43	18	9.9	324	2	BP920932	BP920932 MR2-NT013
44	18	9.9	325	4	BM689940	BM689940 UI-E-CK1-
45	18	9.9	325	5	BU726635	BU726635 UI-E-CK1-

ALIGNMENTS

RESULT 1
LOCUS: CO793602
DEFINITION: NT018B.B1 sc18-22 Neural tube (NT) Ambystoma mexicanum cDNA 5', similar to similar to hypothetical protein MGCS356 (Mus musculus), mRNA sequence.

ACCESSION: CO793602
VERSION: CO793602.1 GI:51009573
KEYWORDS: EST
SOURCE: Ambystoma mexicanum (axolotl)
ORGANISM: Ambystoma mexicanum

REFERENCE: Habermann, B., Bebin, A.G., Herklotz, S., Volkmer, M., Eckelt, K., Pehlke, K., Eppelstein, H.H., Schackert, H.K., Wiebe, G. and Tanaka, B.M. An Ambystoma mexicanum EST sequencing project: Analysis of 17,352 expressed sequence tags from embryonic and regenerating blastema cDNA libraries

JOURNAL: Genome Biol. (2004) in press
COMMENT: Tanaka Lab

Max Planck Institute of Molecular Cell Biology and Genetics, Dresden
Pictenhauerstrasse 108, 01307 Dresden, Germany
Tel: 0049 351 210 2620
Fax: 0049 351 210 1489
Email: tanaka@mpi-cbg.de
Plate: NT018B row: 11 column: B
Seq primer: GCA CAT TAG GCC TAT TTA GGT GAC A.
Location/Qualifiers
1..650
/organism="Ambystoma mexicanum"
/mol_type="mRNA"
/db_xref="taxon:8296"
/tissue_type="Neural Tube, Notochord, Somites"
/cell_type="Includes Neural tube, notochord, somites"
/dev_stage="Stage 18-22"
/clone_id="Sc18-22 Neural tube (NT)"
/note="Vector: pCMVSPORT6; Site 1: NotI; Site 2: SalI; Unnormalized cDNA plasmid library prepared by Invitrogen. Size fractionated mRNA was polyA primed and cloned into NotI-SalI site of pCMVSPORT6. Bacterial host is EMDH108-TONA. Average insert size is 1.5 kb.
TAG_LIB="NT"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 174.066 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-45

Perfect score: 1 GAATTCGACGACTGACGAGC.....TCGCTACCAATGATGATCC 191

Sequence: 1 GAATTCGACGACTGACGAGC.....TCGCTACCAATGATGATCC 191

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

ALIGNMENTS

C	21	18	9.4	3107	9	ADB91252
C	22	18	9.4	3107	10	ADA56256
C	23	18	9.4	4425	3	AACT5665
C	24	18	9.4	4877	10	ADBE5549
C	25	18	9.4	13076	8	ADA41611
C	26	18	9.4	13076	8	ADA41611
C	27	18	9.4	13076	9	ADB91897
C	28	18	9.4	13077	10	ADA57747
C	29	18	9.4	13077	8	ADA41610
C	30	18	9.4	13077	10	ADA57746
C	31	18	9.4	31940	6	AAF28526
C	32	18	9.4	110000	2	AAZ01425
C	33	18	8.9	43	12	ADHA2842
C	34	17	8.9	152	12	ADP94043
C	35	17	8.9	498	12	ADL11516
C	36	17	8.9	1866	13	ADSA47979
C	37	17	8.9	1959	6	AAZ27607
C	38	17	8.9	2394	2	AAZ49305
C	39	17	8.9	2789	6	ABO70842
C	40	17	8.9	2834	8	ABL26522
C	41	17	8.9	4204	8	ADBS3222
C	42	17	8.9	5326	4	ABL10170
C	43	17	8.9	10723	4	ADL14603
C	44	17	8.9	10723	4	ADL14614
C	45	17	8.9	10723	4	ADL14604

Database : N Geneseq 16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	100.0	191	2	AAV23217
2	46	24.1	167	2	AAV23218
3	46	24.1	176	2	AAV23230
4	46	24.1	177	2	AAV23215
5	46	24.1	179	2	AAV23221
6	46	24.1	189	2	AAV23228
7	46	24.1	191	2	AAV23219
8	46	24.1	193	2	AAV23223
9	46	24.1	195	2	AAV23225
10	46	24.1	195	2	AAV23222
11	46	24.1	195	2	AAV23229
12	46	24.1	195	2	AAV23227
13	46	24.1	195	2	AAV23220
14	46	24.1	199	2	AAV23176
15	40	20.9	166	2	AAV23224
16	26	13.6	188	2	AAV23226
17	23	12.0	182	2	AAV23216
18	18	9.4	825	8	ACA38872
19	18	9.4	3107	6	AD33238
20	18	9.4	3107	8	ADA40081

RESULT 1

AAV23217

ID AAV23217 standard; DNA; 191 BP.

AC AAV23217;

DT 28-JUL-1998 (first entry)

DE Saccharomyces cerevisiae promoter Yp13.

XX Saccharomyces cerevisiae; yeast promoter; optimise; spacer;

XX artificial promoter library; gene expression; ds.

OS Synthetic.

OS Saccharomyces cerevisiae.

OS Key

XX Location/Qualifiers

FT 8..181

FT promoter

FT /*tag= a

FT /standard_name= "Yeast promoter"

XX WO9807846-A1.

XX 26-FEB-1998.

XX 25-AUG-1997; 97WO-DK00342.

XX 23-AUG-1996; 96DK-00000886.

XX (JENS// JENSEN P R.

XX Jensen PR, Hammer K;

XX WPI, 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter

XX sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.

XX Claim 28; Page 62; 89pp; English.

XX This is a Saccharomyces cerevisiae promoter sequence used in the

XX construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 934.792 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657b-45

Perfect score: 191

Sequence: 1 GAATTCGTGACTGCTAGGAGC.....TCGCTACCATCATGATCC 191

Scoring table: OLIGO_NUC

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: GenBank1:*

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_lm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	191	100.0	191	6	BD006976	BD006976 Artificia
2	46	24.1	167	6	BD006977	BD006977 Artificia
3	46	24.1	176	6	BD006989	BD006989 Artificia
4	46	24.1	177	6	BD006974	BD006974 Artificia
5	46	24.1	179	6	BD006980	BD006980 Artificia
6	46	24.1	189	6	BD006987	BD006987 Artificia
7	46	24.1	191	6	BD006978	BD006978 Artificia
8	46	24.1	193	6	BD006982	BD006982 Artificia
9	46	24.1	195	6	BD006979	BD006979 Artificia
10	46	24.1	195	6	BD006981	BD006981 Artificia
11	46	24.1	195	6	BD006984	BD006984 Artificia
12	46	24.1	195	6	BD006986	BD006986 Artificia
13	46	24.1	195	6	BD006988	BD006988 Artificia
14	46	24.1	199	6	BD006934	BD006934 Artificia
15	41	21.5	1844	8	SCYOL140W	SCYOL140W
16	41	21.5	9892	8	SCARGCNC	SCARGCNC
17	40	20.9	156	6	BD006983	BD006983 Artificia
18	38	19.9	2479	8	SCYOL141W	SCYOL141W
19	33	17.3	188	6	BD006985	BD006985 Artificia

20	25	13.1	1002	11	CNS06570	AL02010 T7 end of
21	23	12.0	182	6	BD006975	BD006975 Artificia
22	21	11.0	127376	2	AC141672	AC141672 Apis mell
23	21	11.0	196217	2	AC131340	AC131340 Mus muscu
24	20	10.5	143255	2	AC015996	AC015996 Homo sapi
25	20	10.5	153887	9	AC093830	AC093830 Homo sapi
26	20	10.5	199456	9	AC087664	AC087664 Homo sapi
27	20	10.5	237468	2	AC108532	AC108532 Rattus no
28	20	10.5	309638	2	AC118433	AC118433 Rattus no
29	19	9.9	110000	1	AB017308	AB017308 Continuation (5 of
30	19	9.9	200426	9	AC093117	AC093117 Homo sapi
31	19	9.9	214074	2	AC148841	AC148841 Rattus no
32	19	9.9	214455	2	AC118451	AC118451 Rattus no
33	19	9.9	216113	2	AC116121	AC116121 Mus muscu
34	19	9.9	245133	2	AC121033	AC121033 Rattus no
35	19	9.9	246330	2	AC110139	AC110139 Rattus no
36	19	9.9	273684	2	AC113630	AC113630 Rattus no
37	19	9.9	294461	2	AC108628	AC108628 Rattus no
38	18	9.4	4222	9	AB020720	AB020720 Homo sapi
39	18	9.4	4502	9	HSB01483	HSB01483 Homo sapi
40	18	9.4	4955	6	CQ718478	CQ718478 Sequence
41	18	9.4	5082	10	RNCFPR3UT	RNCFPR3UT
42	18	9.4	10155	1	AB012227	AB012227 Xanthomon
43	18	9.4	11371	1	AB006236	AB006236 Pasteurel
44	18	9.4	11951	1	AB010615	AB010615 Fusobacte
45	18	9.4	11951	1	AB010615	AB010615 Fusobacte

ALIGNMENTS

RESULT 1
BD006976
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006976
ACCESSION
BD006976.1 GI:18635347
VERSION
JP 2001503249-A/45.
KEYWORDS
Saccharomyces cerevisiae (baker's yeast)
SOURCE
Saccharomyces cerevisiae
ORGANISM
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE
1 (bases 1 to 191)
AUTHORS
Hammer, K. and Janssen, P. R.
TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL
Patent: JP 2001503249-A 45 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Saccharomyces cerevisiae (yeast)
PN JP 2001503249-A/45
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
FH Topology: Linear;
FT promoter
Location/Qualifiers
(8) .(181).
1. 191
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 191; DB 6; Length 191;
Best local similarity 100.0%; Pred. No. 4e-101;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGTGACTGCTAGGAGTCAGTGGCTCTTGGCGGCGAGGTTGATGCGGC 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1574.46 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-45

Perfect score: 191

Sequence: 1 GAATTCGTGACCTCAGCAGGC.....TCGCTACCAATCATGATCC 191

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	11.0	586	5	BU950278
C 2	20	10.5	586	6	CA844183
C 3	20	10.5	604	4	BM311760
C 4	20	10.5	604	4	BM510073
C 5	20	10.5	616	4	BM351977
C 6	20	10.5	648	6	CA776739
C 7	20	10.5	775	9	CL582491
C 8	20	10.5	1145	6	CB565228
C 9	19	9.9	585	6	CA843478
C 10	19	9.9	588	6	CA776990
C 11	19	9.9	595	6	CA942760
C 12	19	9.9	595	6	CA203485
C 13	19	9.9	616	6	CA233456
C 14	19	9.9	1088	9	CN80577W
C 15	18	9.4	110	6	CA771252
C 16	18	9.4	228	6	CB449507
C 17	18	9.4	235	6	CB440185
C 18	18	9.4	235	6	CB451827
C 19	18	9.4	317	8	BZ671461
C 20	18	9.4	407	8	CC153956
C 21	18	9.4	409	4	BM572543
C 22	18	9.4	409	4	BM572952
C 23	18	9.4	443	7	CR471577
C 24	18	9.4	451	8	CC094891

C 25	18	9.4	452	9	CC842207	CC842207
C 26	18	9.4	469	8	BZ950076	BZ950076
C 27	18	9.4	502	8	BH004907	BH004907
C 28	18	9.4	508	8	BH873913	BH873913
C 29	18	9.4	578	5	BP102832	BP102832
C 30	18	9.4	594	4	BM504542	BM504542
C 31	18	9.4	597	9	CC632433	CC632433
C 32	18	9.4	602	9	CC632440	CC632440
C 33	18	9.4	655	6	CD933432	CD933432
C 34	18	9.4	664	9	BL121165	BL121165
C 35	18	9.4	664	9	CL750452	CL750452
C 36	18	9.4	716	9	BL184773	BL184773
C 37	18	9.4	716	9	CC680315	CC680315
C 38	18	9.4	723	9	BL187483	BL187483
C 39	18	9.4	755	9	BL160384	BL160384
C 40	18	9.4	760	9	CC874751	CC874751
C 41	18	9.4	763	9	BL131577	BL131577
C 42	18	9.4	778	9	CL843043	CL843043
C 43	18	9.4	795	9	CC870420	CC870420
C 44	18	9.4	801	8	CC150321	CC150321
C 45	18	9.4	811	9	CC867413	CC867413

ALIGNMENTS

RESULT 1
BU950278/c
LOCUS
DEFINITION
BU950278.1 GI:24201629
588 bp mRNA linear EST 21-OCT-2002
1079a01.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6132384.3
similar to SW:GLUC_HUMAN P01275 GLUCAGON PRECURSOR. [1] ; mRNA

ACCESSION
BU950278
VERSION
BU950278.1
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Merra, M., Page, D., Wylie, T., Martin, J., Bliscain, A., Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagareishvili, R., Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)

TITLE
JOURNAL
COMMENT
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

FEATURES
source
1. 588
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6132384"
/cruise_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_id="HR85 islet"
/note="Organ: Pancreas; Vector: pluscript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming."

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 152.193 Seconds
(without alignments)

6495.658 Million cell updates/sec

Title: US-09-242-657B-46

Perfect score: 167

Sequence: 1 GAATGCTGATCTCAGGCGCCG.....TCGCTACCATCATGATCC 167

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapept 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167	100.0	167	2	AAV23218 Saccharom
2	76	45.5	177	2	AAV23215 Saccharom
3	76	45.5	195	2	AAV23225 Saccharom
4	76	45.5	195	2	AAV23227 Saccharom
5	76	45.5	195	2	AAV23220 Saccharom
6	76	45.5	199	2	AAV23176 Saccharom
7	68	40.7	189	2	AAV23228 Saccharom
8	68	40.7	195	2	AAV23222 Saccharom
9	65	38.9	193	2	AAV23223 Saccharom
10	58	34.7	179	2	AAV23221 Saccharom
11	51	30.5	176	2	AAV23230 Saccharom
12	51	30.5	195	2	AAV23229 Saccharom
13	48	28.7	191	2	AAV23219 Saccharom
14	46	27.5	191	2	AAV23217 Saccharom
15	40	24.0	166	2	AAV23224 Saccharom
16	39	23.4	188	2	AAV23226 Saccharom
17	23	13.8	182	2	AAV23216 Saccharom
18	18	10.8	1762	10	Ades5903 Rat gene
19	18	10.8	1762	10	Ades5906 Rat gene
20	18	10.8	3432	6	Abz32502 Candida a

ALIGNMENTS

21	18	10.8	5198	6	ABL32278	Ab132278 Human imm
22	18	10.8	6059	6	ABL54343	Ab154343 Chemically
23	18	10.8	110000	2	AAZ01425_08	Continuation (9) of
24	17	10.2	43	12	ADH42842	Adh42842 Novel hum
25	17	10.2	498	12	ADL11516	Adl11516 Cat flea
26	17	10.2	546	12	ACH91742	ACH91742 Human gen
27	17	10.2	570	8	ACA27839	ACA27839 Prokaryot
28	17	10.2	576	6	ABQ44458	Abq44458 Oligonuc1
29	17	10.2	576	6	ABQ44459	Abq44459 Oligonuc1
30	17	10.2	581	12	ACH89633	ACH89633 Human gen
31	17	10.2	907	3	AAZ46787	Aaz46787 Thermococ
32	17	10.2	1866	13	ADG47979	Adg47979 Bacterial
33	17	10.2	1959	6	AAZ27607	Aaz27607 Synechocy
34	17	10.2	2000	6	ABZ15437	Abz15437 Arabidops
35	17	10.2	2000	8	ADA68824	Ada68824 Arabidops
36	17	10.2	2127	12	ADO35448	Ado35448 Novel mou
37	17	10.2	2789	6	ABO70842	Abq70842 Listeria
38	17	10.2	4551	6	ABN85301	Abn85301 Human col
39	17	10.2	4551	6	ABY94763	Abv94763 Human pan
40	17	10.2	5326	4	ABL10170	Ab110170 Drosophila
41	17	10.2	5408	2	AAK77720	Aax77720 Human col
42	17	10.2	5572	6	ABL32815	Ab132815 Human imm
43	17	10.2	5774	13	ADO85964	Adg85964 Human tum
44	17	10.2	5774	13	ACN40861	Acn40861 Tumour-as
45	17	10.2	6000	6	ABL33133	Ab133133 Human imm

RESULT 1
ID AAV23218 standard; DNA; 167 BP.

AAV23218;

28-JUL-1998 (first entry)

Saccharomyces cerevisiae promoter Yp15.

Saccharomyces cerevisiae; yeast promoter; optimise; spacer;

artificial promoter library; gene expression; de.

Synthetic.

Saccharomyces cerevisiae.

Key Location/Qualifiers

promoter 8..167

W09807846-A1.

26-FEB-1998.

25-AUG-1997; 97WO-DK000342.

23-AUG-1996; 96DK-00000886.

(JENS/) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 62; 89pp; English.

This is a Saccharomyces cerevisiae promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 817.331 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-46

Perfect score: 167

Sequence: 1 GAATTCGTACTCAGGCGCGG.....TCGCTACCAATCATGATCC 167

Scoring table: OLIGO_NTC

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank: 1: gb ba: *
2: gb htg: *
3: gb in: *
4: gb om: *
5: gb ov: *
6: gb pat: *
7: gb ph: *
8: gb pl: *
9: gb pr: *
10: gb ro: *
11: gb stb: *
12: gb sy: *
13: gb un: *
14: gb vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	167	100.0	167	6	BD006977
2	76	45.5	177	6	BD006974
3	76	45.5	195	6	BD006979
4	76	45.5	195	6	BD006984
5	76	45.5	195	6	BD006986
6	76	45.5	199	6	BD006934
7	71	42.5	184	8	SCYOL140W
8	71	42.5	9882	8	SCARGC
9	68	40.7	189	6	BD006987
10	68	40.7	195	6	BD006981
11	68	40.7	2479	6	SCYOL141W
12	65	38.9	193	6	BD006982
13	58	34.7	179	6	BD006980
14	51	30.5	176	6	BD006989
15	51	30.5	195	6	BD006988
16	48	28.7	191	6	BD006978
17	46	27.5	191	6	BD006976
18	42	25.1	188	6	BD006985
19	40	24.0	166	6	BD006983

20	25	15.0	1002	11	CNS06570
21	23	13.8	182	6	BD006975
22	21	12.6	127376	2	AC141672
23	21	12.6	196217	2	AC131340
24	20	12.0	59440	2	AC119019_3
25	20	12.0	110000	2	AC107119_1
26	20	12.0	143255	2	AC015996
27	20	12.0	153887	9	AC093830
28	20	12.0	159456	9	AC087664
29	20	12.0	237468	2	AC108532
30	20	12.0	270075	2	AC114451
31	20	12.0	309838	2	AC118433
32	19	11.4	1876	3	AF005529
33	19	11.4	110000	1	AE017308_4
34	19	11.4	110000	2	BX248101_2
35	19	11.4	124185	9	AC007560
36	19	11.4	129675	2	AC102950
37	19	11.4	138006	2	AC068246
38	19	11.4	168738	2	AC093632
39	19	11.4	179503	3	AL445468
40	19	11.4	199814	3	CEB47D3A
41	19	11.4	204900	2	AC120282
42	19	11.4	214074	2	AC148841
43	19	11.4	214455	2	AC118451
44	19	11.4	216113	2	AC116121
45	18	10.8	1561	3	HC0243587

ALIGNMENTS

RESULT 1
LOCUS BD006977 167 bp DNA linear PAT 31-JAN-2002
DEFINITION Artificial promoter libraries for selected organisms and promoters
derived from such libraries.

ACCESSION BD006977.1 GI:18635348

VERSION UP 2001503249-A/46.

KEYWORDS Saccharomyces cerevisiae (baker's yeast)

SOURCE Saccharomyces cerevisiae

ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 167)

AUTHORS Hammer, K. and Janssen, P. R.

TITLE Artificial promoter libraries for selected organisms and promoters
derived from such libraries

JOURNAL Patent: JP 2001503249-A 46 13-MAR-2001;

COMMENT PETER RUDAR JANSSEN

OS Saccharomyces cerevisiae (yeast)

PN JP 2001503249-A/46

PD 13-MAR-2001

PR 25-AUG-1997 JP 1998510287

PI KALIN HAMMER, PETER RUDAR JANSSEN

PC C12N15/09, C12N15/00

CC Strandedness: Double;

CC Topology: Linear;

FT Key Location/Qualifiers

FT promoter (8)..(167).

FEATURES

source

1..167

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 167; DB 6; Length 167;

Best Local Similarity 100.0%; Pred. No. 7, 1e-80;

Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGTACTCAGGCGCGGCTTTCGATGCTATGCGGGGTTTATAAT 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1376.62 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-46

Perfect score: 167

Sequence: 1 GAATTCGTACTACAGGCGCG.....TCGCTACCATCATGATCC 167

Scoring table: OLIGO_NTC

Gapop 60.0 , Gapept 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	12.6	742	9	AG119052 Pan trogl
2	20	12.0	450	2	AM003938 wq84f04.x
3	20	12.0	511	8	AZ977704 2M0253P07
4	20	12.0	775	9	CL582491 OB_Ba004
5	20	12.0	1145	6	CB565228 AGENCOURT
6	19	11.4	258	1	AV011021 AV011021
7	19	11.4	271	1	AV020847 AV020847
8	19	11.4	284	2	BB113023 BB113023
9	19	11.4	285	2	BB113006 BB113006
10	19	11.4	296	2	BB109676 BB109676
11	19	11.4	300	2	BB110358 BB110358
12	19	11.4	301	1	AV018400 AV018400
13	19	11.4	315	1	AV022830 AV022830
14	19	11.4	742	8	BH966805 BH966805
15	18	10.8	228	6	CB449507 703710 MA
16	18	10.8	235	6	CB440185 690440 MA
17	18	10.8	235	6	CB451827 706622 MA
18	18	10.8	301	7	CO025774 EST804158
19	18	10.8	319	1	AV169893 AV169893
20	18	10.8	406	2	AM816055 WK3-ST022
21	18	10.8	407	8	CC153956 CSU-K34.1
22	18	10.8	443	7	CR471577 CR471577
23	18	10.8	451	8	CC094891 CSU-K34.1
24	18	10.8	452	3	CR717693 Tetradon

C	25	18	10.8	452	9	CC842207 ND.L.97010
C	26	18	10.8	471	4	BG511043 sac65b02.
C	27	18	10.8	476	2	AM200801 se93f11.y
C	28	18	10.8	478	2	BF068364 ac90608.y
C	29	18	10.8	493	6	CA581683 EST001358
C	30	18	10.8	494	8	BH1818162 BACPR10-P
C	31	18	10.8	500	7	CO105468 GR_EB003
C	32	18	10.8	502	8	BH004907 BMAC08M2
C	33	18	10.8	519	4	BG405059 sac47e10.
C	34	18	10.8	523	9	CL695644 PR1016d.H
C	35	18	10.8	527	4	BM085813 sa129h04.
C	36	18	10.8	546	2	BH115386 RPCR-24-2
C	37	18	10.8	547	8	BE803625 sr61c12.y
C	38	18	10.8	547	5	BU762070 sa126c01.
C	39	18	10.8	554	5	BQ134085 sa126c01.
C	40	18	10.8	555	4	BI699712 sa126c01.
C	41	18	10.8	563	5	BO296139 sa126c01.
C	42	18	10.8	570	4	BM732101 sa126c01.
C	43	18	10.8	576	4	BI699746 sa126c01.
C	44	18	10.8	578	5	BP102832 BP102832
C	45	18	10.8	589	9	CL674975 PR10113c

ALIGNMENTS

RESULT 1
LOCUS AG119052 742 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-127G02.R, genomic survey sequence.
ACCESSION AG119052
VERSION AG119052.1 GI:16739571
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes

REFERENCE
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 742)

TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
JOURNAL Direct Submission

COMMENT Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpbg@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the Red process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: M13Rev
LIBRARY Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 742
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-127G02.R"
/sex="male"
/cell_type="lymphoblast"
/clone_id="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 12.6% ; Score 21 ; DB 9 ; Length 742 ;
Best Local Similarity 100.0% ; Pred. No. 1.3 ;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 174.066 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-47

Perfect score: 191
Sequence: 1 GAATTCGACTCACCCTC.....TCGCTACCATCATGATCC 191

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	100.0	191	2	AAV23219 Saccharom
2	48	25.1	167	2	AAV23218 Saccharom
3	48	25.1	176	2	AAV23230 Saccharom
4	48	25.1	177	2	AAV23235 Saccharom
5	48	25.1	179	2	AAV23221 Saccharom
6	48	25.1	189	2	AAV23228 Saccharom
7	48	25.1	193	2	AAV23223 Saccharom
8	48	25.1	195	2	AAV23225 Saccharom
9	48	25.1	195	2	AAV23222 Saccharom
10	48	25.1	195	2	AAV23229 Saccharom
11	48	25.1	195	2	AAV23227 Saccharom
12	48	25.1	195	2	AAV23220 Saccharom
13	48	25.1	199	2	AAV23176 Saccharom
14	46	24.1	191	2	AAV23217 Saccharom
15	40	20.9	166	2	AAV23224 Saccharom
16	28	14.7	188	2	AAV23226 Saccharom
17	23	12.0	182	2	AAV23216 Saccharom
18	19	9.9	813	6	ABO31608
19	18	9.4	702	4	AAH84670
20	9.4	702	8	ACA32686	Prokaryot

21	18	9.4	2054	5	AA68385	AA68385 DNA encod
22	18	9.4	2076	5	AA889193	AA889193 DNA encod
23	18	9.4	3202	5	AA885781	AA885781 DNA encod
24	18	9.4	3258	5	AA68365	AA68365 DNA encod
25	18	9.4	6160	6	ABK31273	ABK31273 Signal tr
26	18	9.4	6160	6	AB170234	AB170234 Chemical
27	18	9.4	110000	2	AA201425	Continuation (9 of
28	18	9.4	110000	4	AA199682	Continuation (4 of
29	18	9.4	110000	4	AA199683	Continuation (4 of
30	17	8.9	43	12	ADH42842	ADH42842 Novel hum
31	17	8.9	448	2	AAK84581	AAK84581 G-protein
32	17	8.9	448	3	AAH70647	AAH70647 Human G-P
33	17	8.9	471	9	ACH46842	ACH46842 Human inf
34	17	8.9	498	12	ADL11516	ADL11516 Cat flea
35	17	8.9	801	6	ABQ19916	ABQ19916 Oligonuc
36	17	8.9	801	6	ABQ19917	ABQ19917 Oligonuc
37	17	8.9	1866	13	ADS47979	ADS47979 Bacteri
38	17	8.9	1959	6	AD27607	AD27607 Synchocy
39	17	8.9	2789	6	ABQ70842	ABQ70842 Listeria
40	17	8.9	5326	4	AB110170	AB110170 Drosophi
41	17	8.9	15743	6	ABK28395	ABK28395 DNA trans
42	17	8.9	16026	5	ABH16260	ABH16260 Human ner
43	17	8.9	19115	5	ABA16261	ABA16261 Human ner
44	17	8.9	68355	8	ACF62737	ACF62737 Cancer ba
45	17	8.9	68355	8	ADB20852	ADB20852 MRP1 base

ALIGNMENTS

RESULT 1
AAV23219
ID AAV23219 standard; DNA; 191 BP.
XX
AC AAV23219;
XX
DT 28-JUL-1998 (first entry)
XX
DB Saccharomyces cerevisiae promoter Yp154.
XX
KW Saccharomyces cerevisiae; yeast promoter; optimise; spacer;
KM artificial promoter library; gene expression; ds.
XX
OS Synthetic.
OS Saccharomyces cerevisiae.
XX
FH Key
FH Location/Qualifiers
FT Promoter
FT 8..191
FT /*tag= a
FT /standard_name= "Yeast promoter."
XX
PN WO9807846-A1.
XX
PD 26-FEB-1998.
XX
PP 25-AUG-1997; 97WO-DK000342.
XX
PR 23-AUG-1996; 96DK-00000886.
XX
PA (JENS) JENSEN P R.
XX
PI Jensen PR, Hammer K;
XX
DR WPI, 1998-179062/16.
XX
XX
XX New artificial promoter libraries - containing consensus promoter
XX sequences and variable spacers, used to generate promoters for optimising
XX expression of genes.
XX
XX Claim 28; Page 63; 89pp; English.
XX
XX This is a Saccharomyces cerevisiae promoter sequence used in the
XX construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 934.792 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-47

Perfect score: 191
Sequence: 1 GAATCGTACTGACCGCTC.....TCGCTACCATCATGATCC 191

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_ey:*
12: gb_gy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191	100.0	191	BD006978	BD006978 Artificia
2	48	25.1	167	BD006977	BD006977 Artificia
3	48	25.1	176	BD006989	BD006989 Artificia
4	48	25.1	177	BD006974	BD006974 Artificia
5	48	25.1	179	BD006980	BD006980 Artificia
6	48	25.1	189	BD006987	BD006987 Artificia
7	48	25.1	193	BD006982	BD006982 Artificia
8	48	25.1	195	BD006979	BD006979 Artificia
9	48	25.1	195	BD006981	BD006981 Artificia
10	48	25.1	195	BD006984	BD006984 Artificia
11	48	25.1	195	BD006986	BD006986 Artificia
12	48	25.1	195	BD006988	BD006988 Artificia
13	48	25.1	199	BD006934	BD006934 Artificia
14	46	24.1	191	BD006976	BD006976 Artificia
15	43	22.5	1844	SCYOL140W	SCYOL140W
16	43	22.5	9892	SCARCGDC	SCARCGDC
17	40	20.9	166	BD006983	BD006983 Artificia
18	40	20.9	2479	SCYOL141W	SCYOL141W
19	33	17.3	188	BD006985	BD006985 Artificia

20	30	15.7	1002	11	CNS06JTO	AL402010 T7 end of
21	23	12.0	182	6	BD006975	BD006975 Artificia
22	21	11.0	127376	2	AC141672	AC141672 Apis mell
23	21	11.0	196217	2	AC131340	AC131340 Mus muscu
24	20	10.5	143255	2	AC015996	AC015996 Homo sapi
25	20	10.5	153887	9	AC093830	AC093830 Homo sapi
26	20	10.5	199456	2	AC087664	AC087664 Homo sapi
27	20	10.5	237468	2	AC108532	AC108532 Rattus no
28	20	10.5	309838	2	AC118433	AC118433 Rattus no
29	19	9.9	11226	9	AC109593	AC109593 Homo sapi
30	19	9.9	13385	1	AE009836	AE009836 Pyrobacul
31	19	9.9	110000	19	AE017282_12	Continuation (13 o
32	19	9.9	110000	1	AE017308_4	Continuation (13 o
33	19	9.9	161368	2	AC143686	AC143686 Macaca mu
34	19	9.9	200000	2	AC004624	AC004624 Homo sapi
35	19	9.9	200000	2	AC005141	AC005141 Homo sapi
36	19	9.9	204900	2	AC120282	AC120282 Rattus no
37	19	9.9	214074	2	AC148841	AC148841 Gorilla g
38	19	9.9	214455	2	AC118451	AC118451 Rattus no
39	19	9.9	216113	2	AC116121	AC116121 Mus muscu
40	18	9.4	322	11	BV138705	BV138705 PZA00183
41	18	9.4	373	1	BCRHSFEX3	U5124 Escherichia
42	18	9.4	702	6	AR352357	AR352357 Sequence
43	18	9.4	2853	5	BC081682	BC081682 Dantio rer
44	18	9.4	5082	10	RNCFT3UT	AJ224431 Rattus no
45	18	9.4	6160	6	AX344269	AX344269 Sequence

ALIGNMENTS

RESULT 1
BD006978

LOCUS
BD006978

DEFINITION
Artificial promoter libraries for selected organisms and promoters

ACCESSION
BD006978.1 GI:18635349

VERSION
JP 2001503249-A/47.

KEYWORDS
Saccharomyces cerevisiae (baker's yeast)

SOURCE
Saccharomyces cerevisiae

ORGANISM
Saccharomyces cerevisiae

REFERENCE
1 (bases 1 to 191)

AUTHORS
Hammer, K. and Janssen, P.R.

TITLE
Artificial promoter libraries for selected organisms and promoters

JOURNAL
Patent: JP 2001503249-A 47 13-MAR-2001;

COMMENT
PETER RUDAR JANSSEN

OG Saccharomyces cerevisiae (yeast)

PN JP 2001503249-A/47

PD 13-MAR-2001

PR 25-AUG-1997 JP 1998510287

PI 23-AUG-1996 DK 0886/96

PC KALIN HAMMER, PETER RUDAR JANSSEN

CC C12N15/09, C12N15/00

CC Strandedness: Double;

CC Topology: Linear;

PH Key

FT promoter

Location/Qualifiers

1. 191

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/db_xref="taxon:4932"

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 191; DB 6; Length 191;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATCGTACTGACCGCTCGGTGCAAGGCGGAGATGTGCGGGCGTTCTA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1574.46 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-47

Perfect score: 191

Sequence: 1 GAATGTGTACATCACCGCTC.....TGGCTACCAATCATGATCC 191

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	10.5	775	9	CL582491 OB_Ba004
C 2	20	10.5	1145	6	CB565228 AGENCOURT
C 3	19	9.9	329	9	CG929858 MBEH095TF
C 4	19	9.9	453	1	AI212480 x7d07a1.f
C 5	18	9.4	228	6	CB449507 703710 MA
C 6	18	9.4	235	6	CB440185 690440 MA
C 7	18	9.4	235	6	CB451827 706622 MA
C 8	18	9.4	294	9	CC732676 OGUC018TH
C 9	18	9.4	300	2	BE401196 CNM01P100
C 10	18	9.4	301	6	CD949561 SML_61 Ge
C 11	18	9.4	392	6	CA185827 SCSFST307
C 12	18	9.4	407	8	CC153956 CSU-K34.1
C 13	18	9.4	443	7	CR471577 CR471577
C 14	18	9.4	447	7	CF623831 zmrw05_0
C 15	18	9.4	451	8	CC094891 CSU-K34.1
C 16	18	9.4	452	9	CC842207 ND.L_97010
C 17	18	9.4	454	5	BY255976 BY255976
C 18	18	9.4	476	8	AZ886111 RPCI-23-1
C 19	18	9.4	497	4	BI728379 10310860
C 20	18	9.4	502	8	BH004907 BMBAC08M2
C 21	18	9.4	507	8	AZ886115 RPCI-23-1
C 22	18	9.4	555	2	AW287792 82900809
C 23	18	9.4	577	6	CA192521 SCCST300
C 24	18	9.4	578	5	BP102832 BP102832

C 25	18	9.4	585	9	CL542309 OB_Ba006
C 26	18	9.4	664	9	BX121165 Dantio rer
C 27	18	9.4	681	9	CL750452 OR_BBa011
C 28	18	9.4	687	9	CB571484 tigr-gss
C 29	18	9.4	691	9	AG613302 Escherich
C 30	18	9.4	700	6	CA259393 SCORRT102
C 31	18	9.4	700	6	CD053687 HO03M105
C 32	18	9.4	700	6	CD053936 HO03B025
C 33	18	9.4	705	9	AG113634 Pan trog1
C 34	18	9.4	716	9	CB860315 ND.L_113B1
C 35	18	9.4	723	9	CB860315 ND.L_113B1
C 36	18	9.4	723	9	CB860315 ND.L_113B1
C 37	18	9.4	755	9	CB874751 Dantio rer
C 38	18	9.4	760	9	CB874751 Dantio rer
C 39	18	9.4	763	9	BX131577 Dantio rer
C 40	18	9.4	765	6	CA197505 SCRAPAD102
C 41	18	9.4	778	9	CL843043 OR_CBa007
C 42	18	9.4	790	7	CK703153 ZFI01-P00
C 43	18	9.4	795	8	CC970420 ND.L_12908
C 44	18	9.4	801	8	CC150321 ND.L_60C14
C 45	18	9.4	802	7	CG915902 AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS CL582491/C 775 bp DNA linear GSS 10-JUN-2004
DEFINITION OB_Ba0047119.1 OB_Ba Oryza brachyantha genomic clone
ACCESSION CL582491
VERSION CL582491.1 GI:48540712
KEYWORDS GSS.

ORGANISM

Oryza brachyantha
Oryza brachyantha
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

Kim,H., Yu,Y., Stum,D., Yeast,D., Rao,K., Luo,M., Jetey,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.

TITLE

OMP Project

JOURNAL

Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu

COMMENT

PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 145000 Std Error: 0.00
Plate: 0047 row: 1 column: 19
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES

source location/Qualifiers
1..775
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0047119"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="pH10B"
/clone_lib="OB_Ba"
/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 10.5%; Score 20; DB 9; Length 775;
Best Local Similarity 100.0%; Pred. No. 3.4;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 177.711 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657b-48

Sequence: 1 GAATTCGTGACTCAGAGATTA.....TCGCTACCAATCATGATCC 195

Scoring table: OLIGO_NUC
Gapop_60.0, Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseq11980s:.*
2: geneseq11990s:.*
3: geneseq2000s:.*
4: geneseq2001as:.*
5: geneseq2001bs:.*
6: geneseq2002as:.*
7: geneseq2002bs:.*
8: geneseq2003as:.*
9: geneseq2003bs:.*
10: geneseq2003cs:.*
11: geneseq2003ds:.*
12: geneseq2004as:.*
13: geneseq2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	2	AAV23220 Saccharom
2	76	39.0	167	2	AAV23218 Saccharom
3	76	39.0	177	2	AAV23215 Saccharom
4	76	39.0	195	2	AAV23225 Saccharom
5	76	39.0	195	2	AAV23227 Saccharom
6	76	39.0	199	2	AAV23176 Saccharom
7	68	34.9	189	2	AAV23228 Saccharom
8	68	34.9	195	2	AAV23222 Saccharom
9	65	33.3	193	2	AAV23221 Saccharom
10	58	29.7	179	2	AAV23223 Saccharom
11	51	26.2	176	2	AAV23230 Saccharom
12	51	26.2	195	2	AAV23230 Saccharom
13	48	23.6	191	2	AAV23219 Saccharom
14	46	23.6	191	2	AAV23217 Saccharom
15	40	20.5	166	2	AAV23224 Saccharom
16	39	20.0	188	2	AAV23226 Saccharom
17	23	11.8	182	2	AAV23216 Saccharom
18	20	10.3	344	8	ABX3965 Bovine ES
19	19	9.7	3622	4	ABL18286 Drosophila
20	19	9.7	47670	4	ABL16824 Drosophila

ALIGNMENTS

C	21	19	9.7	50368	4	ABL16768
C	22	18	9.2	3958	2	AAQ29230
C	23	18	9.2	5198	6	ABL32278
C	24	18	9.2	6059	6	ABL54343
C	25	18	9.2	110000	2	AAZ01425_08
C	26	17	8.7	43	12	ADH42842
C	27	17	8.7	356	4	ABZ01398
C	28	17	8.7	498	12	ADL11516
C	29	17	8.7	545	3	AAH78618
C	30	17	8.7	576	6	ABQ44458
C	31	17	8.7	576	6	ABQ44459
C	32	17	8.7	831	12	ADN74244
C	33	17	8.7	883	6	ABQ48934
C	34	17	8.7	883	6	ABQ48935
C	35	17	8.7	899	6	ABQ49044
C	36	17	8.7	899	6	ABQ49045
C	37	17	8.7	967	3	AAA78440
C	38	17	8.7	1721	11	ADM03308
C	39	17	8.7	1721	11	ADSA7979
C	40	17	8.7	1866	13	ADSA7979
C	41	17	8.7	1959	6	ABD27607
C	42	17	8.7	2000	6	ABZ15437
C	43	17	8.7	2000	8	ADA68824
C	44	17	8.7	2064	8	ACA44020
C	45	17	8.7	2064	8	ACA44020

RESULT 1
AAV23220
ID AAV23220 standard; DNA; 195 BP.

XX AAV23220:
XX 28-JUL-1998 (first entry)

XX Saccharomyces cerevisiae promoter Yp18.

XX Saccharomyces cerevisiae; yeast promoter; optimise; spacer;
XX artificial promoter library; gene expression; ds.

OS Synthetic.
OS Saccharomyces cerevisiae.

OS

XX Key Location/Qualifiers

FT promoter /*tag= a

FT /standard_name= "Yeast promoter"

PN WO9807846-A1.

PD 26-FEB-1998.

XX 25-AUG-1997; 97WO-DK000342.

XX 23-AUG-1996; 96DK-00000886.

XX (JENSEN) JENSEN P R.

XX Jensen PR, Hammer K;

XX WPI; 1998-179062/16.

XX New artificial promoter libraries - containing consensus promoter

XX sequences and variable spacers, used to generate promoters for optimising

XX expression of genes.

XX Claim 28; Page 63; 89pp; English.

CC This is a Saccharomyces cerevisiae promoter sequence used in the

CC construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 954.369 Seconds
(without alignments)

9900.540 Million cell updates/sec

Title: US-09-242-657B-48

Perfect score: 195

Sequence: 1 GAATGCTAGCTCAGATTA.....TCGCTACCATCATGATCC 195

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba:*
2: gb_in:*
3: gb_in:*
4: gb_in:*
5: gb_in:*
6: gb_in:*
7: gb_in:*
8: gb_in:*
9: gb_in:*
10: gb_in:*
11: gb_in:*
12: gb_in:*
13: gb_in:*
14: gb_in:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	195	6	BD006979 Artificialia
2	76	39.0	167	6	BD006977 Artificialia
3	76	39.0	177	6	BD006974 Artificialia
4	76	39.0	195	6	BD006984 Artificialia
5	76	39.0	195	6	BD006986 Artificialia
6	76	39.0	199	6	BD006934 Artificialia
7	71	36.4	1844	8	SCYOL140W
8	71	36.4	9892	8	SCARGDC
9	68	34.9	189	6	BD006987 Artificialia
10	68	34.9	195	6	BD006981 Artificialia
11	68	34.9	2479	8	SCYOL141W
12	65	33.3	193	6	BD006982 Artificialia
13	58	29.7	179	6	BD006980 Artificialia
14	51	26.2	176	6	BD006989 Artificialia
15	51	26.2	195	6	BD006988 Artificialia
16	48	24.6	191	6	BD006978 Artificialia
17	46	23.6	191	6	BD006976 Artificialia
18	42	21.5	188	6	BD006985 Artificialia
19	40	20.5	166	6	BD006983 Artificialia

20	12.8	1002	11	CNS06JTO	AL402010 T7 end of
21	11.8	182	6	BD006975	BD006975 Artificialia
22	10.8	127376	2	AC141672	AC141672 Apis mell
23	10.8	196217	2	AC131340	AC131340 Mus muscu
24	10.3	143255	2	AC015996	AC015996 Homo sapi
25	10.3	153887	2	AC093830	AC093830 Homo sapi
26	10.3	199456	9	AC087664	AC087664 Homo sapi
27	10.3	201137	2	AC128312	AC128312 Rattus no
28	10.3	237468	2	AC108332	AC108332 Rattus no
29	10.3	270153	2	AC096262	AC096262 Rattus no
30	10.3	294688	2	AC144498	AC144498 Bos tauri
31	10.3	309838	2	AC118433	AC118433 Rattus no
32	9.7	1876	3	AF300529	AF300529 Diadema
33	9.7	3622	6	CQ596912	CQ596912 Sequence
34	9.7	47670	6	CQ594719	CQ594719 Sequence
35	9.7	50368	6	CQ594635	CQ594635 Sequence
36	9.7	64998	2	AC014035	AC014035 Drosophila
37	9.7	110000	1	AB017308	AB017308 (5 of
38	9.7	112672	2	AC010455	AC010455 Homo sapi
39	9.7	129675	2	AC102950	AC102950 Homo sapi
40	9.7	170195	9	AC025452	AC025452 Homo sapi
41	9.7	176231	3	AC008367	AC008367 Drosophila
42	9.7	179503	9	AL445468	AL445468 Human DNA
43	9.7	188374	3	AC005050	AC005050 Homo sapi
44	9.7	199814	9	CEV47D3A	CEV47D3A Cenorhab
45	9.7	204900	2	AC120282	AC120282 Rattus no

ALIGNMENTS

RESULT 1	BD006979	195 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD006979				
DEFINITION	Artificial promoter libraries for selected organisms and promoters derived from such libraries.				
ACCESSION	BD006979.1	GI:18635350			
VERSION	JP 2001503249-A/48				
KEYWORDS	Saccharomyces cerevisiae (Baker's yeast)				
SOURCE	Saccharomyces cerevisiae				
ORGANISM	Saccharomyces cerevisiae				
REFERENCE	1 (bases 1 to 195)				
AUTHORS	Hammer, K. and Janssen, P.R.				
TITLE	Artificial promoter libraries for selected organisms and promoters derived from such libraries				
JOURNAL	Patent: JP 2001503249-A 48 13-MAR-2001;				
COMMENT	PETER RUDAR JANSSEN				
OS	Saccharomyces cerevisiae (yeast)				
PN	JP 2001503249-A/48				
PD	13-MAR-2001				
PF	25-AUG-1997 JP 1998510287				
PR	23-AUG-1996 DK 0886/96				
PI	KALIN HAMMER, PETER RUDAR JANSSEN				
PC	CI2N15/09, CI2N15/00				
CC	Strandedness: Double;				
CC	Topology: Linear;				
FT	Key				
FEATURES	Location/Qualifiers				
source	1..195				
	/organism="Saccharomyces cerevisiae"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:4932"				
ORIGIN					
Query Match	100.0%	Score 195;	DB 6;	Length 195;	
Best Local Similarity	100.0%	Pred. No. 1.5e-96;			
Matches 195;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
1	GAATGCTAGCTCAGATTAAGCCCGTTGGATTAAGCAGACACTGAGGTGAGAA	60			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1607.43 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-48

Perfect score: 195

Sequence: 1 GAATTCGTCAGTCAGATTA.....TCGCTACCAATCATGATCC 195

Scoring table: OLIGO_NJC

Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	10.3	450	2	AM003938
C 2	20	10.3	775	9	CL582491
C 3	20	10.3	1145	6	CB565228
C 4	19	9.7	519	5	BP061636
C 5	19	9.7	640	5	BP061636
C 6	19	9.7	709	7	CO363631
C 7	19	9.7	735	9	CNS009XU
C 8	19	9.7	742	7	CO367368
C 9	19	9.7	742	8	BH966805
C 10	19	9.7	749	7	CO369242
C 11	19	9.7	874	7	CO369167
C 12	19	9.7	1093	7	CK985339
C 13	18	9.2	228	6	CB449507
C 14	18	9.2	235	6	CB449507
C 15	18	9.2	235	6	CB449507
C 16	18	9.2	240	2	BE817262
C 17	18	9.2	368	7	CF702938
C 18	18	9.2	407	8	CC153956
C 19	18	9.2	422	5	BY001748
C 20	18	9.2	443	7	CR471577
C 21	18	9.2	451	7	CC094891
C 22	18	9.2	452	9	CC842207
C 23	18	9.2	500	4	BH808301
C 24	18	9.2	502	8	BH004907

C 25	18	9.2	521	8	AZ421123
C 26	18	9.2	543	8	AQ119971
C 27	18	9.2	558	8	AQ735481
C 28	18	9.2	578	5	BP102832
C 29	18	9.2	593	8	AZ255847
C 30	18	9.2	623	9	BX388264
C 31	18	9.2	654	5	BM107202
C 32	18	9.2	660	5	BM068182
C 33	18	9.2	664	5	BM121165
C 34	18	9.2	681	9	CL750452
C 35	18	9.2	703	8	CC321410
C 36	18	9.2	716	9	BX184773
C 37	18	9.2	716	9	CC860315
C 38	18	9.2	723	9	BX187483
C 39	18	9.2	750	7	CO560140
C 40	18	9.2	750	8	AZ699940
C 41	18	9.2	755	9	BX160384
C 42	18	9.2	757	9	CG672550
C 43	18	9.2	760	9	CG874751
C 44	18	9.2	763	9	BX131577
C 45	18	9.2	763	9	BX131577

ALIGNMENTS

RESULT 1
AM003938 450 bp mRNA linear EST 08-SEP-1999
LOCUS wg84ff04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2478751 3'
DEFINITION mRNA sequence.

AM003938 1 GI:5850854
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
NATIONAL CANCER INSTITUTE, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Patricia Bernaldo, Ph.D.

CDNA Library: Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrip/image/image.html
Seq primer: -40UP from Gidco.

Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC6"
/note="Vector: pTZ19-D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP GC6 was prepared, and
89 clones were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475597-1476743).
Subtraction by Bento Soares and M. Patricia Bernaldo."

FEATURES

source

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 163.13 Seconds

(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-49

Perfect score: 179
Sequence: 1 GAATCGTGACTCACTAGG.....TCGCTACCAATCATGATCC 179

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	100.0	179	2	AAV23221 Saccharom
2	58	32.4	167	2	AAV23218 Saccharom
3	58	32.4	177	2	AAV23215 Saccharom
4	58	32.4	189	2	AAV23228 Saccharom
5	58	32.4	193	2	AAV23223 Saccharom
6	58	32.4	195	2	AAV23225 Saccharom
7	58	32.4	195	2	AAV23222 Saccharom
8	58	32.4	195	2	AAV23227 Saccharom
9	58	32.4	195	2	AAV23220 Saccharom
10	58	32.4	199	2	AAV23176 Saccharom
11	51	28.5	176	2	AAV23230 Saccharom
12	51	28.5	195	2	AAV23229 Saccharom
13	48	26.8	191	2	AAV23219 Saccharom
14	46	25.7	191	2	AAV23217 Saccharom
15	40	22.3	166	2	AAV23224 Saccharom
16	24	13.4	188	2	AAV23226 Saccharom
17	23	12.8	182	2	AAV23216 Saccharom
18	18	10.1	5198	6	ABL32278 Human imm
19	18	10.1	6059	6	ABL54343 Chemical1
20	18	10.1	101786	3	AAF22293 BAC conta

ALIGNMENTS

21	18	10.1	110000	2	AAZ01425 08	Continuation (9 of
22	17	9.5	43	12	ADH42842	Adh42842 Novel hum
23	17	9.5	429	12	ADP92748	Adp92748 Cotton ex
24	17	9.5	498	12	ADL11516	Adl11516 Cat flea
25	17	9.5	576	6	ABO44458	Abg44458 Oligonuc
26	17	9.5	576	6	ABO44459	Abg44459 Oligonuc
27	17	9.5	1865	13	ADS47979	Ads47979 Bacterial
28	17	9.5	1959	6	ADD27607	Add27607 Synecoc
29	17	9.5	2000	6	ABZ15437	Abz15437 Arabidops
30	17	9.5	2000	8	ADA68824	Ada68824 Arabidops
31	17	9.5	2054	12	ADO62650	Ado62650 Transcript
32	17	9.5	2789	6	ABQ70842	Abq70842 Listeria
33	17	9.5	5326	4	ABL10170	Abi10170 Drosophi
34	17	9.5	5572	6	ABL32815	Abi32815 Human imm
35	17	9.5	6000	6	ABK33133	Abk33133 Human imm
36	17	9.5	6160	6	ABK31273	Abk31273 Signal tr
37	17	9.5	6160	6	ABL70234	Abi70234 Chemical1
38	17	9.5	6163	6	ABN80118	Abn80118 Human che
39	17	9.5	6219	6	ABL32866	Abi32866 Human imm
40	17	9.5	6219	6	AA663324	AA663324 Chemical1
41	17	9.5	6447	6	ABL33086	Abi33086 Human imm
42	17	9.5	6447	6	ABL34532	Abi34532 Human met
43	17	9.5	6447	6	ABL70547	Abi70547 Chemical1
44	17	9.5	6447	6	ABO65973	Abg65973 Human ang
45	17	9.5	6447	7	ADS99793	Ads99793 Bisulphic

RESULT 1	AAV23221	standard; DNA; 179 BP.
ID	AAV23221	standard; DNA; 179 BP.
XX	AAV23221;	
AC	AAV23221;	
XX	28-JUL-1998	(first entry)
DT		
XX	Saccharomyces cerevisiae	promoter Yp183.
DE		
XX	Saccharomyces cerevisiae;	Yeast promoter; Optimise; spacer;
KW	artificial promoter library;	gene expression; ds.
KW		
XX		
OS	Synthetic.	
OS	Saccharomyces cerevisiae.	
XX		
FH	Key	Location/Qualifiers
FT	promoter	8..179
FT		/*tag= a
FT		/standard_name= "Yeast promoter"
XX		
PN	WO9807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENS/)	JENSEN P R.
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI; 1998-179062/16.	
XX		
PT	New artificial promoter libraries -	containing consensus promoter
PT	sequences and variable spacers,	used to generate promoters for optimising
PT	expression of genes.	
XX		
XX	Claim 28; Page 64; 89pp; English.	
XX		
CC	This is a Saccharomyces cerevisiae	promoter sequence used in the
CC	construction of an artificial	promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 876.062 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657b-49

Perfect score: 179
Sequence: 1 GAATTCGAGTACTACTAAGG.....TCGTAACCAATCATGATTC 179

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : GenBank1:

1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	100.0	179	6	BD006980 Artificialia
2	58	32.4	167	6	BD006977 Artificialia
3	58	32.4	177	6	BD006974 Artificialia
4	58	32.4	189	6	BD006987 Artificialia
5	58	32.4	193	6	BD006982 Artificialia
6	58	32.4	195	6	BD006979 Artificialia
7	58	32.4	195	6	BD006981 Artificialia
8	58	32.4	195	6	BD006984 Artificialia
9	58	32.4	195	6	BD006986 Artificialia
10	58	32.4	199	6	BD006934 Artificialia
11	53	29.6	1844	8	SCYOL14OW
12	53	29.6	9892	8	SCARGCDC
13	51	28.5	176	6	BD006989 Artificialia
14	51	28.5	195	6	BD006988 Artificialia
15	50	27.9	2479	8	SCYOL14IW
16	48	26.8	191	6	BD006978 Artificialia
17	46	25.7	191	6	BD006976 Artificialia
18	40	22.3	166	6	BD006983 Artificialia
19	33	18.4	188	6	BD006985 Artificialia

20	23	12.8	182	6	BD006975 Artificialia
21	21	11.7	127376	2	AC141672
22	21	11.7	196217	2	AC133340
23	20	11.2	143255	2	AC015896
24	20	11.2	153887	9	AC093830
25	20	11.2	199456	9	AC087664
26	20	11.2	237468	2	AC108532
27	20	11.2	309838	2	AC118433
28	19	10.6	1876	3	AP300529
29	19	10.6	90161	8	AC066218
30	19	10.6	110000	1	AB017308
31	19	10.6	129675	2	AC102950
32	19	10.6	138218	10	AC132946
33	19	10.6	179503	9	AL445468
34	19	10.6	199814	3	CER47D3A
35	19	10.6	204900	2	AC120282
36	19	10.6	214074	2	AC148841
37	19	10.6	214455	2	AC118451
38	19	10.6	216113	2	AC116121
39	19	10.6	300050	1	AP004595
40	18	10.1	2498	1	BTHKNA
41	18	10.1	5198	6	AX345180
42	18	10.1	6059	6	AX281301
43	18	10.1	11051	1	AE013911
44	18	10.1	11444	1	AE012143
45	18	10.1	11951	1	AB010615

ALIGNMENTS

RESULT 1
BD006980 179 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION
BD006980
VERSION
UP 2001503249-A/49.
KEYWORDS
SOURCE
ORGANISM
Saccharomyces cerevisiae (baker's yeast)
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE
1 (bases 1 to 179)
AUTHORS
Hammer, K. and Janssen, P.R.
TITLE
Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL
Patent: JP 2001503249-A 49 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT

OS Saccharomyces cerevisiae (yeast)
PN JP 2001503249-A/49
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter
Location/Qualifiers
(8) . (179).
1. 179
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 179; DB 6; Length 179;
Best local similarity 100.0%; Pred. No. 3.9e-90;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGAGTACTACTAAGGTTGCCATTACGAATCGCTGTAGAACATCGTAGT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1475.54 Seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-49

Perfect score: 179

Sequence: 1 GAATTCGTAAGTCACTAGG.....TCGCTACCAATCATGATCC 179

Scoring table: OLIGO_NUC

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	11.2	450	2	AM003938
C 2	20	11.2	1145	2	CB565228
C 3	19	10.6	432	2	AM385527
C 4	19	10.6	564	2	AM835565
C 5	19	10.6	742	8	BH966805
C 6	18	10.1	228	6	CB449507
C 7	18	10.1	235	6	CB440185
C 8	18	10.1	235	6	CB451827
C 9	18	10.1	384	1	AI894695
C 10	18	10.1	407	8	CC153956
C 11	18	10.1	451	8	CC094891
C 12	18	10.1	452	9	CC842207
C 13	18	10.1	460	4	BG007453
C 14	18	10.1	483	2	BH069042
C 15	18	10.1	502	8	BH004907
C 16	18	10.1	578	8	BP102832
C 17	18	10.1	593	8	AZ255847
C 18	18	10.1	671	9	CL002843
C 19	18	10.1	681	9	CL750452
C 20	18	10.1	716	9	CC860315
C 21	18	10.1	760	9	CC874751
C 22	18	10.1	778	9	CL843043
C 23	18	10.1	795	9	CC870420
C 24	18	10.1	801	8	CC150321

C 25	18	10.1	811	9	CC867413	CC867413 NDL.1715.
C 26	18	10.1	869	9	CL684606	CL684606 PRI013a.B
C 27	18	10.1	875	9	CL686348	CL686348 PRI014a.B
C 28	18	10.1	879	9	CC851624	CC851624 NDL.12452
C 29	18	10.1	893	9	CC847477	CC847477 NDL.22G12
C 30	18	10.1	913	8	CC068647	CC068647 CSU-K33r.
C 31	18	10.1	923	9	CL467488	CL467488 SAIL.1271
C 32	18	10.1	942	7	CV470564	CV470564 4346T.1 C
C 33	18	10.1	968	8	CC418819	CC418819 PUDIX667D
C 34	18	10.1	984	8	CC109175	CC109175 NDL.51H12
C 35	18	10.1	1008	9	CNS0370X	AL231666 Tetradon
C 36	18	10.1	1037	9	CL058133	CL058133 CH216-871
C 37	18	10.1	1433	8	BZ571137	BZ571137 msh2.1745
C 38	17	9.5	123	8	CC153398	CC153398 CSU-K34.1
C 39	17	9.5	127	8	CC136676	CC136676 NDL.57M12
C 40	17	9.5	128	8	CC134423	CC134423 NDL.77M19
C 41	17	9.5	139	8	CC142414	CC142414 NDL.70116
C 42	17	9.5	145	9	CC864019	CC864019 NDL.101P1
C 43	17	9.5	152	8	CC087331	CC087331 CSU-K33r.
C 44	17	9.5	152	8	CC153578	CC153578 CSU-K34.1
C 45	17	9.5	157	5	BP746873	BP746873 BP746873

ALIGNMENTS

RESULT 1
AM003938/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AM003938 450 bp mRNA linear EST 08-SEP-1999
wg84f04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2478751 3',
mRNA sequence.
AM003938.1 GI:5850854
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLM at:
www.bio.lnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES

source

1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_id="NCI_CGAP_GC6"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalised library NCI_CGAP GC4 was prepared, and
as circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 177.711 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-50

Sequence: 1 GAATGCTAGCTAGATATCC.....TCGCTACCAATCATGATCC 195

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn180bs:*
2: geneseqn190bs:*
3: geneseqn200bs:*
4: geneseqn200bs:*
5: geneseqn200bs:*
6: geneseqn200bs:*
7: geneseqn200bs:*
8: geneseqn200bs:*
9: geneseqn200bs:*
10: geneseqn200bs:*
11: geneseqn200bs:*
12: geneseqn200bs:*
13: geneseqn200bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	2	AAV23222 Saccharom
2	76	39.0	189	2	AAV23228 Saccharom
3	68	34.9	157	2	AAV23218 Saccharom
4	68	34.9	177	2	AAV23215 Saccharom
5	68	34.9	195	2	AAV23225 Saccharom
6	68	34.9	195	2	AAV23227 Saccharom
7	68	34.9	195	2	AAV23220 Saccharom
8	68	34.9	199	2	AAV23176 Saccharom
9	65	33.3	193	2	AAV23223 Saccharom
10	58	29.7	179	2	AAV23221 Saccharom
11	51	26.2	176	2	AAV23220 Saccharom
12	51	26.2	195	2	AAV23229 Saccharom
13	48	24.6	191	2	AAV23219 Saccharom
14	45	23.6	191	2	AAV23217 Saccharom
15	40	20.5	166	2	AAV23224 Saccharom
16	34	17.4	188	2	AAV23226 Saccharom
17	23	11.8	182	2	AAV23216 Saccharom
18	18	9.2	5198	6	ABL54343 Human imm
19	18	9.2	6059	6	ABL54343 Chemical1
20	18	9.2	110000	2	AAZ01425_08 Continuation (9 of

ALIGNMENTS

21	17	8.7	43	12	ADH42842	Adh42842 Novel hum
22	17	8.7	60	3	AA65331	AA65331 Angiotens
23	17	8.7	60	6	ABN49863	ABN49863 Human BPI
24	17	8.7	60	8	AB272701	AB272701 Detection
25	17	8.7	498	12	ADL11516	ADL11516 Cat flea
26	17	8.7	576	6	ABQ44458	ABQ44458 Oligonuc1
27	17	8.7	576	6	ABQ44459	ABQ44459 Oligonuc1
28	17	8.7	944	6	ABK52155	ABK52155 Haemophil
29	17	8.7	1866	13	AD847979	AD847979 Bacterial
30	17	8.7	1959	6	AD27607	AD27607 Synchocy
31	17	8.7	2000	6	AB215437	AB215437 Arabidops
32	17	8.7	2000	6	ADA68824	ADA68824 Bacteri
33	17	8.7	2181	10	ADF03385	ADF03385 Bacteri
34	17	8.7	2789	8	AB070842	AB070842 Listeria
35	17	8.7	5049	8	ABX63327	ABX63327 Human CDN
36	17	8.7	5049	10	ADB47467	ADB47467 Human CDN
37	17	8.7	5093	12	ADN05625	ADN05625 Antipsoi
38	17	8.7	5326	4	AB110170	AB110170 Drosophi
39	17	8.7	5572	6	AB132815	AB132815 Human imm
40	17	8.7	6000	6	AB133133	AB133133 Human imm
41	17	8.7	6160	6	ABK31273	ABK31273 Signal tr
42	17	8.7	6160	6	ABU70234	ABU70234 Chemical1
43	17	8.7	6163	6	ABN80118	ABN80118 Human che
44	17	8.7	6219	6	AB132866	AB132866 Human imm
45	17	8.7	6219	6	AA633324	AA633324 Chemical1

RESULT 1	AAV23222	standard; DNA; 195 BP.
ID	AAV23222	standard; DNA; 195 BP.
XX	AAV23222;	
AC	28-JUL-1998	(first entry)
XX		
DT		
XX		
DE	Saccharomyces cerevisiae promoter Yp190.	
XX		
KW	Saccharomyces cerevisiae; yeast promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; ds.	
OS	Synthetic.	
OS	Saccharomyces cerevisiae.	
XX		
PH	Key	Location/Qualifiers
FT	Promoter	8..190
FT		/*tag= a
FT		/standard_name= "Yeast promoter"
XX		
PN	W09807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENS)/ JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI; 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX		
PS	Claim 28; Page 64-65; 89pp; English.	
XX		
CC	This is a Saccharomyces cerevisiae promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 954.369 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-50

Perfect score: 195

Sequence: 1 GAATTCGACGATCAGTATCC.....TCGCTACCAATCATGATCC 195

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBml:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	6	BD006981 Artificia
2	76	39.0	189	6	BD006987 Artificia
3	68	34.9	167	6	BD006977 Artificia
4	68	34.9	177	6	BD006974 Artificia
5	68	34.9	195	6	BD006979 Artificia
6	68	34.9	195	6	BD006984 Artificia
7	68	34.9	195	6	BD006986 Artificia
8	68	34.9	199	6	BD006934 Artificia
9	65	33.3	193	6	BD006982 Artificia
10	63	32.3	1844	8	SCYOL140W
11	63	32.3	9852	8	SCARGCDC
12	60	30.8	2479	8	SCYOL141W
13	58	29.7	179	6	BD006980 Artificia
14	51	26.2	176	6	BD006989 Artificia
15	51	26.2	195	6	BD006988 Artificia
16	48	24.6	191	6	BD006978 Artificia
17	46	23.6	191	6	BD006976 Artificia
18	40	20.5	166	6	BD006983 Artificia
19	34	17.4	188	6	BD006985 Artificia

20	C	21	11.8	182	6	BD006975	Artificia
21	C	21	10.8	127376	2	AC141672	Apis mell
22	C	21	10.8	196217	2	AC131340	Mus muscu
23	C	20	10.3	143255	2	AC015996	AC015996 Homo sapi
24	C	20	10.3	153887	9	AC093830	AC093830 Homo sapi
25	C	20	10.3	173414	9	AC010387	AC010387 Homo sapi
26	C	20	10.3	179875	10	BX004998	BX004998 Mouse DNA
27	C	20	10.3	185531	2	AC012059	AC012059 Homo sapi
28	C	20	10.3	199456	6	AC087664	AC087664 Homo sapi
29	C	20	10.3	237468	2	AC108532	AC108532 Rattus no
30	C	20	10.3	309838	2	AC118433	AC118433 Rattus no
31	C	19	9.7	1876	3	AF300529	AF300529 Diadema
32	C	19	9.7	110000	1	AE017308	Continuation (5 of
33	C	19	9.7	129675	2	AC102950	AC102950 Homo sapi
34	C	19	9.7	160344	2	AC130227	AC130227 Rattus no
35	C	19	9.7	175965	10	AC121921	AC121921 Mus muscu
36	C	19	9.7	177995	9	AC016700	AC016700 Homo sapi
37	C	19	9.7	179503	9	AL445468	AL445468 Human DNA
38	C	19	9.7	199814	3	CEX47D3A	CEX47D3A Caenorhab
39	C	19	9.7	204800	2	AC120282	AC120282 Rattus no
40	C	19	9.7	209401	2	AC108575	AC108575 Rattus no
41	C	19	9.7	214074	2	AC148841	AC148841 Gorilla g
42	C	19	9.7	214455	2	AC118451	AC118451 Rattus no
43	C	19	9.7	216113	2	AC116121	AC116121 Mus muscu
44	C	18	9.2	2498	1	BTHKNA	U03552 Bacillus th
45	C	18	9.2	5198	6	AX345180	AX345180 Sequence

ALIGNMENTS

RESULT 1
BD006981
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION
BD006981.1 GI:18635352
VERSION
UP 2001503249-A/50.
KEYWORDS
Saccharomyces cerevisiae (baker's yeast)
SOURCE
Saccharomyces cerevisiae
ORGANISM
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetes.
REFERENCE
Hammer, K. and Janssen, P.R.
AUTHORS
Artificial promoter libraries for selected organisms and promoters
TITLE
derived from such libraries
JOURNAL
Patent: UP 2001503249-A 50 13-MAR-2001;
PETER RUDAR JANSSEN
COMMENT
OS Saccharomyces cerevisiae (yeast)
PN UP 2001503249-A/50
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PF 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter
Location/Qualifiers
(8). (190).

FEATURES

source

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 195; DB 6; Length 195;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGACGATCAGTATCCAGGCTGTTGAGGCTGTCGAGGTTAGCAGCGGCGGG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1607.43 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-50

Perfect score: 195
Sequence: 1 GAATGCTGACATGATATCC.....TCGCTACCATATGATATCC 195

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	10.3	450	2	AM003938
C 2	20	10.3	649	9	CR211414
C 3	20	10.3	1145	6	CB565228
C 4	19	9.7	442	7	CN962453
C 5	19	9.7	742	8	BH966805
C 6	19	9.7	813	6	B249606
C 7	19	9.2	228	6	CB449507
C 8	18	9.2	235	6	CB440185
C 9	18	9.2	235	6	CB451827
C 10	18	9.2	309	8	B2884389
C 11	18	9.2	362	4	BM157710
C 12	18	9.2	407	8	CC153956
C 13	18	9.2	421	8	B2932856
C 14	18	9.2	451	8	CC094891
C 15	18	9.2	452	9	CC842207
C 16	18	9.2	502	8	BH004907
C 17	18	9.2	503	8	AO508743
C 18	18	9.2	539	8	AL924323
C 19	18	9.2	574	8	AO508759
C 20	18	9.2	575	4	BM866093
C 21	18	9.2	578	4	BP102832
C 22	18	9.2	584	4	BM863594
C 23	18	9.2	593	8	AZ255847
C 24	18	9.2	654	5	BM107202

C 25	18	9.2	660	5	BM068182
C 26	18	9.2	681	9	CL750452
C 27	18	9.2	683	5	BM107288
C 28	18	9.2	685	7	CF399811
C 29	18	9.2	716	9	CC860315
C 30	18	9.2	725	7	CF471681
C 31	18	9.2	760	9	CC874751
C 32	18	9.2	768	7	CF471818
C 33	18	9.2	778	9	CL843043
C 34	18	9.2	795	9	CC870420
C 35	18	9.2	801	8	CC150321
C 36	18	9.2	804	6	CB696606
C 37	18	9.2	811	9	CC867413
C 38	18	9.2	845	7	CF399803
C 39	18	9.2	869	9	CL684606
C 40	18	9.2	879	9	CC851624
C 41	18	9.2	893	9	CC847477
C 42	18	9.2	913	8	CC068647
C 43	18	9.2	952	9	CNS06P15
C 44	18	9.2	958	8	CC109175
C 45	18	9.2	1008	9	CNS0370X

ALIGNMENTS

RESULT 1
AM003938/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AM003938 450 bp mRNA linear EST 08-SEP-1999
wg84f04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2478751 3',
mRNA sequence.
AM003938.1 GI:5850854
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 450)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
www.bio.lnlnl.gov/bbrp/image/image.html
Seq primer: -40bp from Gibco.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI_CGAP GC6"
/note="Vector: pTT3-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP GC4 was prepared, and
as circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subcloning hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES

source

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 175.888 Seconds

(Without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657b-51

Perfect score: 193
Sequence: 1 GAATTCGTCATCATGCTG.....TCGTCATCATGATGATTC 193

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001as:*
5: geneeqn2001bs:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	193	100.0	193	2	AAV23223 Saccharom
2	65	33.7	167	2	AAV23218 Saccharom
3	65	33.7	177	2	AAV23215 Saccharom
4	65	33.7	189	2	AAV23228 Saccharom
5	65	33.7	195	2	AAV23225 Saccharom
6	65	33.7	195	2	AAV23222 Saccharom
7	65	33.7	195	2	AAV23227 Saccharom
8	65	33.7	195	2	AAV23220 Saccharom
9	65	33.7	199	2	AAV23176 Saccharom
10	58	30.1	179	2	AAV23221 Saccharom
11	51	26.4	176	2	AAV23230 Saccharom
12	51	26.4	195	2	AAV23229 Saccharom
13	48	24.9	191	2	AAV23219 Saccharom
14	46	23.8	191	2	AAV23217 Saccharom
15	40	20.7	166	2	AAV23224 Saccharom
16	31	16.1	188	2	AAV23226 Saccharom
17	23	11.9	182	2	AAV23216 Saccharom
18	19	9.8	432	5	AA572548 DNA encod
19	19	9.8	432	5	AA570405 DNA encod
20	19	9.8	500	6	AB571881 Human GTP

ALIGNMENTS

c	21	19	9.8	2825	13	ADT05412	Adt05412 Haemophil
c	22	19	9.8	85814	13	ADT05644	Adt05644 Haemophil
c	23	19	9.8	110000	2	AA742063_05	Continuation (6 of
c	24	18	9.3	479	5	ABV42910	ABV42910 Human pro
c	25	18	9.3	479	5	ABV34043	ABV34043 Human pro
c	26	18	9.3	581	5	ABV12917	ABV12917 Human pro
c	27	18	9.3	604	5	ABV57870	ABV57870 Human pro
c	28	18	9.3	2307	10	ABZ77159	ABZ77159 Human pro
c	29	18	9.3	5198	6	ABL32278	ABL32278 Human imm
c	30	18	9.3	6059	6	ABL54343	ABL54343 Chemical
c	31	18	9.3	110000	2	AAZ01425_08	Continuation (9 of
c	32	18	9.3	110000	6	ABO74964_2	Continuation (3 of
c	33	17	8.8	25	9	ACT19082	Act19082 Human mic
c	34	17	8.8	43	12	ADH42842	Adh42842 Novel hum
c	35	17	8.8	121	12	ADK33845	Adk33845 Polynucle
c	36	17	8.8	252	8	ABZ56801	Abz56801 Aspergill
c	37	17	8.8	498	12	ADL11516	Adl11516 Cat flea
c	38	17	8.8	576	6	ABO44458	Abg44458 Oligonuc
c	39	17	8.8	576	6	ABO44459	Abg44459 Oligonuc
c	40	17	8.8	840	2	AA785619	AA785619 PD498 alk
c	41	17	8.8	840	2	AAV47626	AAV47626 Nucleotid
c	42	17	8.8	840	2	AAV45120	AAV45120 Bacillus
c	43	17	8.8	840	2	AAZ07552	AAZ07552 Bacillus
c	44	17	8.8	840	3	AAZ61246	AAZ61246 DNA encod
c	45	17	8.8	840	3	AAA30782	AAA30782 Bacillus

RESULT 1	AAV23223	standard; DNA; 193 BP.
ID	AAV23223	standard; DNA; 193 BP.
XX	AAV23223;	(first entry)
AC	AAV23223;	(first entry)
XX	28-JUL-1998	(first entry)
DT	28-JUL-1998	(first entry)
XX	Saccharomyces cerevisiae	promoter Yp191.
DE	Saccharomyces cerevisiae	Yeast promoter; optimise; spacer;
XX	Saccharomyces cerevisiae	artificial promoter library; gene expression; ds.
KW	Saccharomyces cerevisiae	artificial promoter library; gene expression; ds.
XX	Saccharomyces cerevisiae	artificial promoter library; gene expression; ds.
OS	Saccharomyces cerevisiae	artificial promoter library; gene expression; ds.
XX	Saccharomyces cerevisiae	artificial promoter library; gene expression; ds.
FT	Key	Location/Qualifiers
FT	promoter	8.189
FT	/*tag= a	/standard_name= "Yeast promoter"
XX	WO9807846-A1.	
XX	26-FEB-1998.	
PD	26-FEB-1998.	
XX	25-AUG-1997;	97WO-DK000342.
PF	25-AUG-1997;	97WO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
PR	23-AUG-1996;	96DK-00000886.
XX	(JENS)/ JENSEN P R.	
PA	(JENS)/ JENSEN P R.	
XX	Jensen PR, Hammer K;	
PI	Jensen PR, Hammer K;	
XX	WPI; 1998-179062/16.	
DR	WPI; 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
PT	expression of genes.	
XX	Claim 28; Page 65; 89pp; English.	
XX	This is a Saccharomyces cerevisiae promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 944.58 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657b-51

Sequence: 1 GAATTCGTGACTCATGCTG.....TCGCTACCAATCATGATGCC 193

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	193	100.0	193	6	BD006982
2	65	33.7	167	6	BD006977
3	65	33.7	177	6	BD006974
4	65	33.7	189	6	BD006987
5	65	33.7	195	6	BD006979
6	65	33.7	195	6	BD006981
7	65	33.7	195	6	BD006984
8	65	33.7	195	6	BD006986
9	65	33.7	199	6	BD006982
10	60	31.1	1844	6	BD006934
11	60	31.1	9892	8	SCARGCPC
12	58	30.1	179	6	BD006980
13	57	29.5	2479	8	SCYOL141W
14	51	26.4	176	6	BD006989
15	51	26.4	195	6	BD006988
16	48	24.8	191	6	BD006978
17	46	23.8	191	6	BD006976
18	40	20.7	166	6	BD006983
19	33	17.1	188	6	BD006985

20	c	21	11.9	182	6	BD006975	BD006975	Artificia
21	c	21	10.9	127376	2	AC141672	AC141672	Apis mell
22	c	21	10.9	196217	2	AC131340	AC131340	Mus muscu
23	c	23	10.4	135855	9	HS135814	AC015996	Homo sapi
24	c	25	10.4	143255	2	AC015996	AC015996	Homo sapi
25	c	25	10.4	153887	9	AC093830	AC093830	Homo sapi
26	c	26	10.4	161121	9	AF235093	AF235093	Homo sapi
27	c	27	10.4	199456	9	AC087664	AC087664	Homo sapi
28	c	28	10.4	237468	2	AC108532	AC108532	Homo sapi
29	c	30	10.4	309838	2	AC118433	AC118433	Rattus no
30	c	30	10.4	340000	9	HS21C010	HS21C010	Homo sapi
31	c	31	9.8	1876	3	AF300529	AF300529	Didascia
32	c	32	9.8	2825	6	CO872835	CO872835	Sequence
33	c	33	9.8	10372	1	U32734	U32734	Haemophilu
34	c	34	9.8	33434	2	AC138536	AC138536	Homo sapi
35	c	35	9.8	37806	2	AC138493	AC138493	Homo sapi
36	c	36	9.8	39832	2	AC138500	AC138500	Homo sapi
37	c	37	9.8	40913	2	AC138501	AC138501	Homo sapi
38	c	38	9.8	56999	2	AC008876	AC008876	Homo sapi
39	c	39	9.8	85814	6	CO873067	CO873067	Sequence
40	c	40	9.8	93479	9	AF201337	AF201337	Homo sapi
41	c	41	9.8	110000	1	AE017308	AE017308	Continuation (5 of
42	c	42	9.8	110000	6	BD426631	BD426631	Continuation (6 of
43	c	43	9.8	110000	6	AR274513	AR274513	Continuation (6 of
44	c	44	9.8	110000	6	AB541453	AB541453	Continuation (6 of
45	c	45	9.8	120099	9	AC011449	AC011449	Homo sapi

ALIGNMENTS

RESULT 1
BD006982 193 bp DNA linear PAT 31-JAN-2002
LOCUS Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
BD006982
ACCESSION BD006982.1 GI:18635353
VERSION JP 2001503249-A/51.
KEYWORDS Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 193)
AUTHORS Hammer, K. and Janssen, P.R.
TITLE Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURNAL Patent: JP 2001503249-A 51 13-MAR-2001;
PETER RUDAR JANSSEN
OS Saccharomycetes cerevisiae (yeast)
PN JP 2001503249-A/51
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (8)..(189).
Location/Qualifiers
1..193
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 193; DB 6; Length 193;
Best Local Similarity 100.0%; Pred. No. 7.9e-89;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAATTCGTGACTCATGCTGCGGCGGCGAGGTGTACTTCCCATTTGAGTGA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1590.95 Seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-51

Perfect score: 193
Sequence: 1 GAATTCGTAACATCATGCTG.....TCGCTACCAATCATGATGCC 193

Scoring table: Oligo_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hic:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	10.4	327	5	BP149593
C 2	20	10.4	450	2	AM003938
C 3	20	10.4	880	9	CG143758
C 4	20	10.4	976	9	CG188502
C 5	20	10.4	1145	6	CB555228
C 6	19	9.8	385	1	AA856829
C 7	19	9.8	432	1	AA828154
C 8	19	9.8	482	8	AZ640410
C 9	19	9.8	507	4	BH286542
C 10	19	9.8	515	4	BH286542
C 11	19	9.8	551	9	BH286542
C 12	19	9.8	629	2	AM979129
C 13	19	9.8	662	8	BH306835
C 14	19	9.8	674	8	AZ572687
C 15	19	9.8	676	5	BM446905
C 16	19	9.8	690	1	AV860723
C 17	19	9.8	690	5	BM267840
C 18	19	9.8	712	5	BM143396
C 19	19	9.8	735	5	BM157660
C 20	19	9.8	742	8	BH966805
C 21	19	9.8	789	5	BM137989
C 22	19	9.8	789	5	BM094899
C 23	18	9.3	112	7	CN188567
C 24	18	9.3	143	9	CC625845

25	18	9.3	228	6	CB449507	CB449507
C 26	18	9.3	235	6	CB440185	CB440185
C 27	18	9.3	235	6	CB451827	CB451827
C 28	18	9.3	407	8	CC153956	CC153956
C 29	18	9.3	451	8	CC094891	CC094891
C 30	18	9.3	452	9	CC842207	CC842207
C 31	18	9.3	502	8	BH004907	BH004907
C 32	18	9.3	512	5	BH084906	BH084906
C 33	18	9.3	578	5	BP102832	BP102832
C 34	18	9.3	589	9	CE430166	CE430166
C 35	18	9.3	593	8	AZ255847	AZ255847
C 36	18	9.3	639	2	BM457437	BM457437
C 37	18	9.3	639	2	BM457437	BM457437
C 38	18	9.3	639	5	BM262095	BM262095
C 39	18	9.3	654	5	BM107202	BM107202
C 40	18	9.3	654	5	BM454462	BM454462
C 41	18	9.3	660	5	BM068182	BM068182
C 42	18	9.3	660	8	AZ572235	AZ572235
C 43	18	9.3	664	5	BM342053	BM342053
C 44	18	9.3	668	8	AQ488589	AQ488589
C 45	18	9.3	671	9	AG170493	AG170493

ALIGNMENTS

RESULT 1
LOCUS BP149593/c 327 bp mRNA linear EST 30-DEC-2003
DEFINITION BP149593 full-length enriched swine cDNA library, adult ovary
Accession BP149593.1 GI:40399066
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa (pig)
Sus scrofa
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 327)
Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
Okumura,N., Hamasima,N. and Awata,T.
PEDE (Pig EST Data Explorer): construction of a database for ESTs
derived from porcine full-length cDNA libraries
Nucleic Acids Res. 32 (1), D484-D488 (2004)
CONTACT: Hirohide Uenishi
National Institute of Agrobiological Sciences
National Institute of Agrobiological Sciences
2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
Tel: +81-29-838-8627
Fax: +81-29-838-8627
Email: huenishi@affrc.go.jp
EST project with full-length enriched cDNA libraries carried out in
Animal Genome Research Program (Japan) by National Institute of
Agrobiological Sciences and STAFF-Institute
Single pass sequencing of clones derived from oligo-capped cDNA
library
Vector sequences were eliminated by RepeatMasker version 2002/07/13
and crossmatch version 0.900319
Low quality bases were trimmed based on the quality values.

FEATURES

source

1..327
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="OVRM10136D09"
/tissue_type="Ovary"
/dev_stage="adult"
/clone_lib="full-length enriched swine cDNA library, adult
ovary"

ORIGIN

Query Match 10.4%; Score 20; DB 5; Length 327;
Best Local Similarity 100.0%; Pred. No. 7.4;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 151.282 Seconds

(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657b-52

Perfect score: 166

Sequence: 1 GAATCGTGACTCATTAG.....TCGTAACATCATGATCC 166

Scoring table: OLIGO_NUC
Gapop 60.0 ; Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003s:*
7: geneseqn2004s:*
8: geneseqn2005s:*
9: geneseqn2006s:*
10: geneseqn2007s:*
11: geneseqn2008s:*
12: geneseqn2009s:*
13: geneseqn2010s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	100.0	166	2	AAV23224 Saccharom
2	40	24.1	167	2	AAV23218 Saccharom
3	40	24.1	176	2	AAV23230 Saccharom
4	40	24.1	177	2	AAV23215 Saccharom
5	40	24.1	179	2	AAV23221 Saccharom
6	40	24.1	189	2	AAV23228 Saccharom
7	40	24.1	191	2	AAV23217 Saccharom
8	40	24.1	191	2	AAV23219 Saccharom
9	40	24.1	193	2	AAV23223 Saccharom
10	40	24.1	195	2	AAV23225 Saccharom
11	40	24.1	195	2	AAV23222 Saccharom
12	40	24.1	195	2	AAV23229 Saccharom
13	40	24.1	195	2	AAV23227 Saccharom
14	40	24.1	195	2	AAV23220 Saccharom
15	40	24.1	199	2	AAV23176 Saccharom
16	23	13.9	182	2	AAV23216 Saccharom
17	21	12.7	2829	10	ADG37055 Saccharom
18	19	11.4	1280	4	ABL28585 Drosophila
19	19	11.4	5310	4	ABL28584 Drosophila
20	18	10.8	483	9	ACH28410 Human adu

ALIGNMENTS

RESULT 1	AAV23224	standard; DNA, 166 BP.
ID	AAV23224	standard; DNA, 166 BP.
XX	AAV23224	
AC	AAV23224	
DT	28-JUN-1998	(first entry)
XX		
DE	Saccharomyces cerevisiae promoter Yp192.	
XX		
KW	Saccharomyces cerevisiae; yeast promoter; optimise; spacer;	
RV	artificial promoter library; gene expression; ds.	
XX		
OS	Synthetic.	
OS	Saccharomyces cerevisiae.	
XX		
FT	Key	Location/Qualifiers
FT	promoter	8..166
FT		/*tag= a
FT		/strand_name= "Yeast promoter"
XX		
FN	W09807846-A1.	
XX		
PD	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97NO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENS/) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI; 1998-179062/16.	
XX		
FT	New artificial promoter libraries - containing consensus promoter	
FT	sequences and variable spacers, used to generate promoters for optimising	
FT	expression of genes.	
XX		
PS	Claim 28; Page 66; 89pp; English.	
XX		
CC	This is a Saccharomyces cerevisiae promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

21	18	10.8	484	6	ABN21204	Abn21204 Human ORF
22	18	10.8	986	6	ABL89468	ABL89468 Human pol
23	18	10.8	1812	12	ADG97999	Adg97999 Human can
24	18	10.8	2320	4	ABL04798	ABL04798 Drosophila
25	18	10.8	4333	3	AAAI5173	AAAI5173 DNA encod
26	18	10.8	5326	4	ABL10170	ABL10170 Drosophila
27	18	10.8	17062	4	ABL14164	ABL14164 Drosophila
28	18	10.8	22475	12	ADG97998	Adg97998 Human can
29	18	10.8	39451	13	ABD33424	ABD33424 Human can
30	18	10.8	110000	2	AAZ01425_08	Continuation. (9 of
31	18	10.8	110000	11	ACN45090_0	Actn45090 Human gen
32	18	10.8	191150	12	ADM69029	Adm69029 Human pla
33	17	10.2	43	12	ADH42842	Adh42842 Novel hum
34	17	10.2	924	13	ADG47946	Adg47946 Bacterial
35	17	10.2	1322	11	ACH96046	Ach96046 Klebsiella
36	17	10.2	1395	8	ACA28429	Acta28429 Prokaryot
37	17	10.2	1866	13	ADG47979	Adg47979 Bacterial
38	17	10.2	1959	6	AAD27607	Aad27607 Synchococ
39	17	10.2	2000	8	ADA71611	Ada71611 Rice gene
40	17	10.2	2789	6	ABQ70842	Abq70842 Listeria
41	17	10.2	2934	4	ABL18212	ABL18212 Drosophila
42	17	10.2	4263	4	ABL18878	ABL18878 Drosophila
43	17	10.2	5423	6	ABV78084	Abv78084 Hypoxia-r
44	17	10.2	5423	12	ADQ20972	Adq20972 Human sof
45	17	10.2	7892	6	ABK40055	Abk40055 Human che

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: June 21, 2005, 12:37:20 ; Search time 812.437 Seconds

(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-52

Perfect score: 166
Sequence: 1 GAATTCGACTCACTTAAAG.....TCGCTCAATCATGATCC 166

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	100.0	166	6	BD006983 Artificialia
2	40	24.1	167	6	BD006977 Artificialia
3	40	24.1	176	6	BD006989 Artificialia
4	40	24.1	177	6	BD006974 Artificialia
5	40	24.1	179	6	BD006980 Artificialia
6	40	24.1	189	6	BD006987 Artificialia
7	40	24.1	191	6	BD006976 Artificialia
8	40	24.1	191	6	BD006978 Artificialia
9	40	24.1	193	6	BD006982 Artificialia
10	40	24.1	193	6	BD006979 Artificialia
11	40	24.1	195	6	BD006981 Artificialia
12	40	24.1	195	6	BD006984 Artificialia
13	40	24.1	195	6	BD006986 Artificialia
14	40	24.1	195	6	BD006988 Artificialia
15	40	24.1	199	6	BD006934 Artificialia
16	35	21.1	1844	8	SCYOL140W
17	35	21.1	9892	8	SCARGCDC
18	33	19.9	188	6	BD006985 Artificialia
19	32	19.3	2479	8	SCYOL141W

20	23	13.9	182	6	BD006975 Artificialia
21	22	13.3	199456	9	AC087664 Homo sapi
22	21	12.7	2517	1	BS049358 Bacillus su
23	21	12.7	2829	1	AB095984 Bacillus
24	21	12.7	44560	9	AC004536 Homo sapi
25	21	12.7	138277	8	AP003273 Oryza sat
26	21	12.7	169562	8	AP003381 Oryza sat
27	21	12.7	196217	2	AC131340 Mus muscu
28	21	12.7	207829	1	BS0B0010 Bacillus su
29	21	12.7	314838	3	CERY105C5A Caenorhab
30	20	12.0	88982	5	AL732598 Zebrfish
31	20	12.0	127376	2	AC141672 Apis mell
32	20	12.0	189787	2	CR293534 Dario rer
33	20	12.0	237468	2	AC108532 Rattus no
34	20	12.0	309838	2	AC118433 Rattus no
35	19	11.4	1280	6	C0612360 Sequence
36	19	11.4	1950	3	AF536215 Drosophi
37	19	11.4	2718	3	BT001588 Drosophi
38	19	11.4	5310	6	C0612359 Sequence
39	19	11.4	14720	5	AB092690 Oryzias l
40	19	11.4	18221	2	AC017207 Drosophi
41	19	11.4	110000	2	AC073804 Mus muscu
42	19	11.4	118008	2	AC149793 Aedes aeg
43	19	11.4	133176	2	AC138445 Felis cat
44	19	11.4	143255	2	AC015996 Homo sapi
45	19	11.4	146464	5	BX530024 Zebrfish

ALIGNMENTS

RESULT 1
BD006983 166 bp DNA linear PAT 31-JAN-2002
LOCUS
DEFINITION Artificial promoter libraries for selected organisms and promoters derived from such libraries.
ACCESSION BD006983
VERSION BD006983.1 GI:18635354
KEYWORDS UP 2001503249-A/52.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 166)
AUTHORS Hammer,K. and Janssen,P.R.
TITLE Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL Patent: JP 2001503249-A 52 13-MAR-2001;
PETER RUDAR JANSSEN
COMMENT OS Saccharomyces cerevisiae (yeast)
PN JP 2001503249-A/52
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PF 23-AUG-1996 DK 0866/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC CL2N15/09, CL2N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (8)..(166).
Location/Qualifiers
1..166
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

FEATURES

source

ORIGIN

Query Match 100.0%; Score 166; DB 6; Length 166;
Best Local Similarity 100.0%; Pred. No. 2,1e-84;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGACTCACTTAAAGCTACGAGATTAGATCAAGTCGGAATTAATTTA 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1368.38 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-52

Sequence: 1 GAATCGTACGCTACCTTAG.....TCGCTACCAATCATGATCC 166

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	13.3	1013	9	CG687610 ZMMBC016
2	21	12.7	611	9	CG257777 OGSB1677C
3	20	12.0	1597	4	BG420056 602453743
4	19	11.4	407	8	CC153956 CSU-K34.1
5	19	11.4	497	9	CL625841 OR_BBA002
6	19	11.4	573	9	CL727351 OR_BBA005
7	19	11.4	585	8	AZ304626 IM0004113
8	19	11.4	629	9	CL711249 OR_BBA003
9	19	11.4	653	9	CC683056 ND1.120C8
10	19	11.4	668	4	CL749538 OR_BBA011
11	19	11.4	716	9	BI510143 BB170028A
12	19	11.4	726	9	CC680315 ND1.113E1
13	19	11.4	730	8	CC108649 ND1.7F8.S
14	19	11.4	737	8	CC073396 CSU-K33r.
15	19	11.4	760	9	CC874751 ND1.36A22
16	19	11.4	798	8	CC125920 ND1.5H24
17	19	11.4	818	8	CC152619 CSU-K34.1
18	19	11.4	866	8	CC526010 EY08364.D
19	19	11.4	894	8	CC080755 CSU-K33r.
20	19	11.4	907	8	CC074727 CSU-K33r.
21	19	11.4	913	8	CC068647 CSU-K33r.
22	19	11.4	983	8	CC068079 CSU-K33r.
23	19	11.4	1567	8	CC305966 CH261-173
24	19	11.4			

C 25	18	10.8	123	8	CC153398
C 26	18	10.8	127	8	CC136676
C 27	18	10.8	128	8	CC134423
C 28	18	10.8	139	8	CC142414
C 29	18	10.8	145	9	CC664019
C 30	18	10.8	152	8	CC087331
C 31	18	10.8	171	8	CC138744
C 32	18	10.8	197	8	CC116691
C 33	18	10.8	171	8	CC126051
C 34	18	10.8	208	9	CC651377
C 35	18	10.8	220	8	CC123851
C 36	18	10.8	223	8	CC085442
C 37	18	10.8	228	6	CB449507
C 38	18	10.8	229	8	CC121245
C 39	18	10.8	231	8	CC139977
C 40	18	10.8	232	8	CC117932
C 41	18	10.8	235	6	CB440185
C 42	18	10.8	235	6	CB451827
C 43	18	10.8	249	9	CC653848
C 44	18	10.8	251	8	CC075473
C 45	18	10.8	251	8	CC141557

ALIGNMENTS

RESULT 1
LOCUS CG687610 1013 bp DNA linear GSS 08-OCT-2003
DEFINITION ZMMBC0165C20f ZMMBC (EcoRI) Zea mays genomic clone ZMMBC0165C20
5', genomic survey sequence.
ACCESSION CG687610 GI:37585000
VERSION CG687610.1 GI:37585000
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 1013)
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zochovec,V., Fuks,G., Yu,X., Wing,R. and Messing,J.
Sequencing of the maize genome at PgrR (2003c)
Unpublished (2003)
Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seg primer: T7
Classes: BAC ends
High quality sequence start: 114.
Location/Qualifiers
1..1013
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/CD_xref="taxon:4577"
/clone="ZMMBC0165C20"
/lab_host="E. coli DH10B"
/clone_1lb="ZMMBC (EcoRI)"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN
Query Match 13.3%; Score 22; DB 9; Length 1013;
Best local similarity 100.0%; Pred. No. 0.34;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
34 TTTAGATCTAAGTCGGAATA 55
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 177.711 Seconds

(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-53

Perfect score: 195

Sequence: 1 GAATTCGTACTACATGCGCC.....TCGCTACCATCATGATGCC 195

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	195	2	AAV23225 Saccharom
2	76	39.0	167	2	AAV23218 Saccharom
3	76	39.0	177	2	AAV23215 Saccharom
4	76	39.0	195	2	AAV23227 Saccharom
5	76	39.0	195	2	AAV23220 Saccharom
6	76	39.0	199	2	AAV23176 Saccharom
7	68	34.9	189	2	AAV23228 Saccharom
8	68	34.9	195	2	AAV23222 Saccharom
9	65	33.3	193	2	AAV23223 Saccharom
10	58	29.7	179	2	AAV23221 Saccharom
11	51	26.2	176	2	AAV23230 Saccharom
12	51	26.2	195	2	AAV23229 Saccharom
13	48	24.6	191	2	AAV23219 Saccharom
14	46	23.6	191	2	AAV23217 Saccharom
15	40	20.5	166	2	AAV23224 Saccharom
16	39	20.0	188	2	AAV23226 Saccharom
17	23	11.8	182	2	AAV23216 Saccharom
18	19	9.7	816	6	ABQ44081 Oligonuc
19	19	9.7	816	6	ABQ44080 Oligonuc
20	18	9.2	5198	6	ABL32278 Human imm

ALIGNMENTS

21	18	9.2	6059	6	ABL54343	Ab154343 Chemical
22	18	9.2	110000	2	AAZ01425	Continuation (9 of
23	17	8.7	43	12	ADH42842	Adh42842 Novel hum
24	17	8.7	498	12	ADL11516	Adl11516 Cat Flea
25	17	8.7	557	6	ABQ23603	Abq23603 Oligonuc
26	17	8.7	557	6	ABQ23602	Abq23602 Oligonuc
27	17	8.7	576	6	ABQ44458	Abq44458 Oligonuc
28	17	8.7	576	6	ABQ44459	Abq44459 Oligonuc
29	17	8.7	1012	10	ADC29924	Adc29924 Human nov
30	17	8.7	1866	13	ADS47979	AdS47979 Bacterial
31	17	8.7	1959	6	AD27607	Ad27607 Synchocy
32	17	8.7	2000	8	AB215437	Ab215437 Arabidops
33	17	8.7	2000	8	ADA68824	Ada68824 Arabidops
34	17	8.7	2789	6	ABQ70842	AbQ70842 Listeria
35	17	8.7	2940	8	ABT33956	AbT33956 Human pig
36	17	8.7	4071	4	AAU05562	AAU05562 Human rep
37	17	8.7	5326	4	AB110170	Ab110170 Drosophila
38	17	8.7	5572	6	AB132815	Ab132815 Human imm
39	17	8.7	6000	6	AB133133	Ab133133 Human imm
40	17	8.7	6160	6	ABK31273	AbK31273 Signal tr
41	17	8.7	6160	6	AB170234	Ab170234 Chemical
42	17	8.7	6163	6	ABN80118	Abn80118 Human che
43	17	8.7	6219	6	AB132866	Ab132866 Human imm
44	17	8.7	6219	6	AA563324	AA563324 Chemical
45	17	8.7	6447	6	ABL33086	Ab133086 Human imm

RESULT 1

ID AAV23225 standard; DNA; 195 BP.

AAV23225;

28-JUL-1998 (first entry)

Saccharomyces cerevisiae promoter Yp212.

Saccharomyces cerevisiae; yeast promoter; optimise; spacer;

artificial promoter library; gene expression; ds.

Saccharomyces cerevisiae.

Key Location/Qualifiers

promoter 8..190 /*tag= a

W09807846-A1. /standard_name= "Yeast promoter."

26-FEB-1998.

25-AUG-1997; 97MO-DX000342.

23-AUG-1996; 96DK-00000886.

(JENSEN) JENSEN P R.

Jensen PR, Hammer K;

WPI, 1998-179062/16.

New artificial promoter libraries - containing consensus promoter

sequences and variable spacers, used to generate promoters for optimising

expression of genes.

Claim 28; Page 66; 89pp; English.

This is a Saccharomyces cerevisiae promoter sequence used in the

construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 954.369 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-53

Perfect score: 195
Sequence: 1 GAATTCGTACTGACGTCGCC.....TCGCTACCAATCATGATTC 195

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	195	100.0	195	6	BD006984 Artificia
2	75	39.0	157	6	BD006977 Artificia
3	76	39.0	177	6	BD006974 Artificia
4	76	39.0	195	6	BD006979 Artificia
5	76	39.0	195	6	BD006986 Artificia
6	76	39.0	199	6	BD006934 Artificia
7	75	38.5	1844	8	SCYOL140W Artificia
8	75	38.5	9892	8	SCARGCDC Artificia
9	72	36.9	2479	8	SCYOL141W Artificia
10	68	34.9	189	6	BD006987 Artificia
11	68	34.9	195	6	BD006981 Artificia
12	68	33.3	193	6	BD006982 Artificia
13	58	29.7	179	6	BD006980 Artificia
14	51	26.2	176	6	BD006989 Artificia
15	51	26.2	195	6	BD006988 Artificia
16	48	24.6	191	6	BD006978 Artificia
17	46	23.6	191	6	BD006976 Artificia
18	42	21.5	188	6	BD006985 Artificia
19	40	20.5	166	6	BD006983 Artificia

20	27	13.8	1002	11	CNS06JTO
21	23	11.8	182	6	BD006975
22	22	11.8	249878	2	AC117848
23	21	10.8	127376	2	AC141672
24	21	10.8	196217	2	AC131340
25	20	10.3	143255	2	AC015996
26	20	10.3	153887	9	AC093830
27	20	10.3	199456	9	AC087664
28	20	10.3	237468	2	AC108532
29	20	10.3	309838	3	AC118433
30	19	9.7	1876	3	AF300529
31	19	9.7	4624	2	AC017957
32	19	9.7	110000	1	AE017308
33	19	9.7	110000	2	AC128384
34	19	9.7	129675	2	AC102950
35	19	9.7	176284	2	AC108444
36	19	9.7	179503	9	AL445468
37	19	9.7	180684	2	AC133098
38	19	9.7	199679	10	AC131695
39	19	9.7	199814	3	CEY47D3A
40	19	9.7	204900	2	AC120282
41	19	9.7	214074	2	AC148841
42	19	9.7	214455	2	AC118451
43	19	9.7	216113	2	AC116121
44	19	9.7	223650	2	AC095283
45	19	9.7	270387	2	AC128402

ALIGNMENTS

RESULT 1
BD006984 195 bp DNA linear PAT 31-JAN-2002
LOCUS Artificial promoter libraries for selected organisms and promoters
DEFINITION derived from such libraries.

ACCESSION BD006984 GI:18635355
VERSION UP 2001503249-A/53.
KEYWORDS Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM *Saccharomyces cerevisiae*
Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

REFERENCE 1 (bases 1 to 195)
AUTHORS Hammer, K. and Janssen, P.R.

TITLE Artificial promoter libraries for selected organisms and promoters
JOURNAL derived from such libraries
PATENT: JP 2001503249-A 53 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT OS Saccharomyces cerevisiae (yeast)
PN UP 2001503249-A/53
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN

PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT promoter (8)..(190).
Location/Qualifiers

FEATURES
Source 1..195
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 195; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.3e-103; Indels 0; Gaps 0;
Matches 195; Conservative 0; Mismatches 0;

1 GAATTCGTACTGACGTCGCCGACATGATGATGATTTAAACCCGATTAATAC 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 171.332 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657b-54

Perfect score: 188
Sequence: 1 GAATTCGTTACTCAGACAGC.....TCGCTACCAATCATGATTC 188

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	63.3	188	2	AAV23226 Saccharom
2	43	22.9	177	2	AAV23215 Saccharom
3	43	22.9	195	2	AAV23227 Saccharom
4	42	22.3	167	2	AAV23218 Saccharom
5	42	22.3	195	2	AAV23225 Saccharom
6	42	22.3	195	2	AAV23220 Saccharom
7	42	22.3	195	2	AAV23176 Saccharom
8	34	18.1	189	2	AAV23228 Saccharom
9	34	18.1	195	2	AAV23222 Saccharom
10	33	17.6	166	2	AAV23224 Saccharom
11	33	17.6	176	2	AAV23230 Saccharom
12	33	17.6	179	2	AAV23221 Saccharom
13	33	17.6	191	2	AAV23217 Saccharom
14	33	17.6	191	2	AAV23219 Saccharom
15	33	17.6	193	2	AAV23223 Saccharom
16	33	17.6	195	2	AAV23229 Saccharom
17	23	12.2	182	2	AAV23216 Saccharom
18	18	9.6	5928	4	ABL10555 Drosophil
19	18	9.6	9285	4	ABL10554 Drosophil
20	18	9.6	110000	2	AAZ01425_08 Continuation (9 of

21	17	9.0	43	12	ADH42842 Novel hum
22	17	9.0	201	13	ADQ47658 Myocardia
23	17	9.0	618	13	ADQ48728 Novel can
24	17	9.0	842	4	AAH04161 Human CDN
25	17	9.0	1866	13	ADSA47979 Bacterial
26	17	9.0	1959	6	AAAD27607 Synchocy
27	17	9.0	2000	6	ABE215437 Arabidops
28	17	9.0	2000	8	ADA68824 Arabidops
29	17	9.0	2213	4	AB122542 Drosophil
30	17	9.0	2432	4	AAH17226 Human CDN
31	17	9.0	3451	10	ADD18495 Human PRO
32	17	9.0	3453	11	AD131997 Human CDN
33	17	9.0	3453	11	ADN95184 Human BEC
34	17	9.0	3453	13	ADR87633 Human Int
35	17	9.0	3853	6	AB192113 Human Tum
36	17	9.0	3853	10	ABX72038 DNA encod
37	17	9.0	4507	6	ABQ95501 Human cod
38	17	9.0	6059	6	ABU54343 Chemical
39	17	9.0	6408	13	ACN38219 Tumour-as
40	17	9.0	7512	9	ACH03907 Human CDN
41	17	9.0	7512	10	AD156430 Chicken c
42	17	9.0	8305	6	ABL33568 Human imm
43	17	9.0	10724	13	ADQ38670 Human SNP
44	17	9.0	10872	6	ABK31389 Signal tr
45	17	9.0	10872	6	ABL70336 Chemical

ALIGNMENTS

RESULT 1	AAV23226	standard; DNA; 188 BP.
ID	AAV23226	standard; DNA; 188 BP.
XX	AAV23226;	(first entry)
XX	28-JUL-1998	(first entry)
XX	Saccharomyces cerevisiae promoter Yp24.	
XX	Saccharomyces cerevisiae; yeast promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; ds.	
KW	Synthetic.	
OS	Saccharomyces cerevisiae.	
XX	Key	Location/Qualifiers
FT	promoter	8..183
FT		/*tag= a
FT		/standard_name= "Yeast promoter"
XX	WO9807846-A1.	
XX	26-FEB-1998.	
XX	25-AUG-1997;	97WO-DK000342.
XX	23-AUG-1996;	96DK-00000886.
XX	(JENS/) JENSEN P R.	
XX	Jensen PR, Hammer K;	
XX	WPI; 1998-179062/16.	
XX	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX	Claim 28; Page 67; 89pp; English.	
XX	This is a Saccharomyces cerevisiae promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 920.109 Seconds

(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-54

Perfect score: 188
Sequence: 1 GAATTCGACTACAGACAGC.....TCGCTACCAATCATGATCC 188

Scoring table: OLIGO_NTC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_str: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	100.0	188	6	BD006985
2	143	22.9	177	6	BD006974
3	43	22.9	195	6	BD006986
4	42	22.3	167	6	BD006977
5	42	22.3	195	6	BD006979
6	42	22.3	195	6	BD006984
7	42	22.3	199	6	BD006934
8	42	22.3	184	8	SCYOL140W
9	42	22.3	2479	8	SCYOL141W
10	42	22.3	9892	8	SCARGCDC
11	34	18.1	189	6	BD006987
12	34	18.1	195	6	BD006981
13	33	17.6	166	6	BD006983
14	33	17.6	176	6	BD006989
15	33	17.6	179	6	BD006980
16	33	17.6	191	6	BD006976
17	33	17.6	191	6	BD006978
18	33	17.6	193	6	BD006982
19	33	17.6	195	6	BD006988

20	25	13.3	1002	11	CNS06JTO
21	23	12.2	182	6	BD006975
22	20	10.6	1094	5	AY543825
23	20	10.6	222162	5	AC147885
24	19	10.1	1584	8	AF397019
25	19	10.1	2814	8	AF395759
26	19	10.1	129675	2	AC102950
27	19	10.1	135332	2	AP064863
28	19	10.1	155086	8	AC113249
29	19	10.1	175584	9	PTB080A15
30	19	10.1	179503	9	AL445468
31	19	10.1	182740	9	BS000163
32	19	10.1	190319	9	BS000158
33	19	10.1	204174	2	CR318615
34	19	10.1	214098	9	AC097381
35	19	10.1	340000	9	HS21C083
36	18	9.6	1703	8	AF395758
37	18	9.6	1853	3	AY636149
38	18	9.6	2498	1	BTHKNA
39	18	9.6	4107	1	RME313328
40	18	9.6	5082	10	RNCFT307
41	18	9.6	5928	6	CQ585315
42	18	9.6	8890	1	SVATPS
43	18	9.6	9285	6	CQ585314
44	18	9.6	9971	1	AB000820
45	18	9.6	11051	1	AB013911

ALIGNMENTS

RESULT 1
BD006985
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS
PN
PD
PF
PR
PI
PC
CC
FH
FT

BD006985
Artificial promoter libraries for selected organisms and promoters derived from such libraries.
BD006985
JP 2001503249-A/54.
Saccharomyces cerevisiae (baker's yeast)
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
1 (bases 1 to 188)
Hammer, K. and Janssen, P.R.
Artificial promoter libraries for selected organisms and promoters derived from such libraries
Patent: JP 2001503249-A 54 13-MAR-2001;
PETER RUDAR JANSSEN

OS
PN
PD
PF
PR
PI
PC
CC
FH
FT

OS
PN
PD
PF
PR
PI
PC
CC
FH
FT

OS
PN
PD
PF
PR
PI
PC
CC
FH
FT

ORIGIN

Query Match 100.0%; Score 188; DB 6; Length 188;

Best Local Similarity 100.0%; Pred. No. 7.1e-95; Indels 0; Gaps 0;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGACTACAGACAGCCTTATGACTTCAGAGACGCTACTTCGGTGGCTACTA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using gw model

Run on: June 21, 2005, 20:41:42 ; Search time 1549.73 Seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-54

Sequence: 1 GAATTCGTAAGTCAACAGCAGC.....TCGTAACCAATCATGATCC 188

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	11.2	576	9	CE047678 tigr-gss-
2	20	10.6	447	6	CA710148 wdk2c.pko
3	20	10.6	450	2	AM003938 wq84f04.x
4	20	10.6	775	9	CL582491 OB_Ba004
5	19	10.1	440	7	CF840673 pshB009xw
6	19	10.1	476	7	CF848824 p8MA008xg
7	19	10.1	521	9	CL596753 OB_Ba005
8	19	10.1	532	7	CF614609 CES008893
9	19	10.1	646	7	CF858063 pamy0081B
10	19	10.1	696	7	BE583367 8-BB-MY P
11	19	10.1	726	7	CK107475 SB02043B2
12	19	10.1	736	8	CE446923 tigr-gss-
13	19	10.1	742	8	BZ658474 OGCA025TM
14	19	10.1	742	8	BH966805 od126d05.
15	19	10.1	953	8	CC329047 OGPA017TV
16	18	9.6	332	8	BH199821 sm1-57L2.
17	18	9.6	344	8	BH879324 h445g03.b
18	18	9.6	443	7	CR471577 CR471577
19	18	9.6	443	1	AA129151 zn29402.r
20	18	9.6	469	8	BZ950076 CH240_39C
21	18	9.6	528	9	BH004907 BMBAC08M2
22	18	9.6	528	9	CC695603 OGMS30TV
23	18	9.6	530	7	CO785646 BL284A.F0
24	18	9.6	530	9	CL565572 OB_Ba003

25	18	9.6	549	8	BZ708535
26	18	9.6	617	8	AZ69802
27	18	9.6	654	8	BW107202
28	18	9.6	659	8	BZ714067
29	18	9.6	660	5	BW068182
30	18	9.6	664	9	BX121165
31	18	9.6	683	5	BW107288
32	18	9.6	689	8	BZ969667
33	18	9.6	716	9	BX184773
34	18	9.6	723	9	BX187483
35	18	9.6	741	5	BW345848
36	18	9.6	755	9	BX160384
37	18	9.6	763	9	BX131577
38	18	9.6	777	8	BZ714058
39	18	9.6	817	8	AZ535637
40	18	9.6	831	9	CC500494
41	18	9.6	832	7	CK603478
42	18	9.6	846	5	BW182465
43	18	9.6	1008	9	CMS0370X
44	18	9.6	1028	5	BW851529
45	18	9.6	1529	7	CK401902

ALIGNMENTS

RESULT 1
CE047678 576 bp DNA linear GSS 24-SEP-2003
tigr-gss-dog-17000357487772 Dog Library Canis familiaris genomic,
LOCUS
DEFINITION
genomic survey sequence.

CE047678
CE047678.1 GI:35086580
GSS.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness, E.F., Batina, V., Halpern, A.L., Levy, S., Remington, K.,
Ruch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shotgun.
Location/Qualifiers
1..576
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: Bact1; Libraries were prepared from
peripheral blood"

FEATURES
SOURCE

ORIGIN

Query Match 11.2%, Score 21; DB 9; Length 576;
Best local similarity 100.0%; Pred. No. 1.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

141 TTTTTCGTTTCAATTAAT 161
DB 201 TTTTTCGTTTCAATTAAT 221

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 177.711 Seconds
(without alignment)
6495.658 Million cell updates/sec

Title: US-09-242-657B-55

Sequence: 1 GAATTCGTCGATCAATGAT.....TCGCTACCAATCAGATCC 195

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database : N_Geneseq_15Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	2	AAV23227 Saccharom
2	77	39.5	177	2	AAV23215 Saccharom
3	76	39.0	167	2	AAV23218 Saccharom
4	76	39.0	195	2	AAV23225 Saccharom
5	76	39.0	195	2	AAV23220 Saccharom
6	76	39.0	199	2	AAV23176 Saccharom
7	68	34.9	189	2	AAV23228 Saccharom
8	68	34.9	195	2	AAV23222 Saccharom
9	65	33.3	193	2	AAV23223 Saccharom
10	58	29.7	179	2	AAV23221 Saccharom
11	51	26.2	176	2	AAV23230 Saccharom
12	51	26.2	195	2	AAV23229 Saccharom
13	48	24.6	191	2	AAV23219 Saccharom
14	46	23.6	191	2	AAV23217 Saccharom
15	40	20.5	166	2	AAV23224 Saccharom
16	39	20.0	188	2	AAV23226 Saccharom
17	23	11.8	182	2	AAV23216 Saccharom
18	18	9.2	410	5	AAFe6626 Novel hum
19	18	9.2	1485	11	ACN89791 Breast ca
20	18	9.2	1738	6	ABX17322 Human can

c	21	18	9.2	2392	12	ADK00588	Adk00588 HOMO prot
c	22	18	9.2	5198	6	ABL32278	ABL32278 Human imm
c	23	18	9.2	6059	6	ABL54343	ABL54343 Chemical
c	24	18	9.2	12683	5	AAI98939	AAI98939 Human exc
c	25	18	9.2	12683	5	AAI63289	AAI63289 Human exc
c	26	18	9.2	12744	4	AAI98938	AAI98938 Human exc
c	27	18	9.2	12744	5	AAI63288	AAI63288 Human kid
c	28	18	9.2	110000	2	AAZ01425	Continuation (9 of
c	29	17	8.7	43	12	ADH42842	Adh42842 Novel hum
c	30	17	8.7	169	4	AAI26112	AAI26112 Probe #16
c	31	17	8.7	169	4	ABA73453	ABA73453 Human foe
c	32	17	8.7	169	4	AAI53887	AAI53887 Probe #22
c	33	17	8.7	169	4	ABA38770	ABA38770 Probe #17
c	34	17	8.7	169	4	AAK48059	AAK48059 Human bon
c	35	17	8.7	169	4	AAK21896	AAK21896 Human bra
c	36	17	8.7	169	4	ABE47773	ABE47773 Human liv
c	37	17	8.7	169	6	ABS21990	ABS21990 Human gen
c	38	17	8.7	169	12	ACH85646	ACH85646 Human gen
c	39	17	8.7	498	12	ADL11516	Adl11516 Cat flea
c	40	17	8.7	576	6	ABQ44458	ABQ44458 Oligonuc
c	41	17	8.7	576	6	ABQ44459	ABQ44459 Oligonuc
c	42	17	8.7	591	4	AAI16907	AAI16907 Probe #68
c	43	17	8.7	591	4	ABA60876	ABA60876 Human foe
c	44	17	8.7	591	4	AAI40771	AAI40771 Probe #94
c	45	17	8.7	591	4	ABA28866	ABA28866 Probe #73

ALIGNMENTS

RESULT 1	AAV23227	standard; DNA; 195 BP.
ID	AAV23227	
XX	AAV23227;	
AC	28-JUL-1998	(first entry)
XX		
DE	Saccharomyces cerevisiae promoter Yp435.	
XX		
KW	Saccharomyces cerevisiae; Yeast promoter; optimise; spacer;	
XX	artificial promoter library; gene expression; ds.	
OS	Saccharomyces cerevisiae.	
XX		
FT	Key	Location/Qualifiers
FT	promoter	8..190
FT		/*tag= a
FT		/standard_name= "Yeast promoter"
XX		
XX	MO9807846-A1.	
XX		
XX	26-FEB-1998.	
XX		
XX	25-AUG-1997;	97WO-DK000342.
XX		
XX	23-AUG-1996;	96DK-00000886.
XX		
XX	(JENSEN/ JENSEN P R.	
XX		
XX	Jensen PR, Hammer K;	
XX		
XX	WPI, 1998-179062/16.	
XX		
XX	New artificial promoter libraries - containing consensus promoter	
XX	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX		
XX	Claim 28; Page 67; 89pp; English.	
XX		
XX	This is a Saccharomyces cerevisiae promoter sequence used in the	
XX	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 954.369 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657b-55

Perfect score: 1 GAATTCGTACTAATGAT.....TCGCTACCATCATGATCC 195

Sequence: 1 GAATTCGTACTAATGAT.....TCGCTACCATCATGATCC 195

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: GenBank:

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	6	BD006986 Artificialia
2	177	39.5	177	6	BD006974 Artificialia
3	76	39.0	167	6	BD006977 Artificialia
4	76	39.0	195	6	BD006979 Artificialia
5	76	39.0	195	6	BD006984 Artificialia
6	76	39.0	199	6	BD006934 Artificialia
7	71	36.4	1844	8	SCYOL140W
8	71	36.4	9892	8	SCARGCDD
9	68	34.9	189	6	BD006987 Artificialia
10	68	34.9	195	6	BD006981 Artificialia
11	65	33.3	2479	6	SCYOL141W
12	65	33.3	193	6	BD006982 Artificialia
13	58	29.7	179	6	BD006980 Artificialia
14	51	26.2	176	6	BD006989 Artificialia
15	51	26.2	195	6	BD006988 Artificialia
16	48	24.6	191	6	BD006978 Artificialia
17	48	24.6	191	6	BD006976 Artificialia
18	43	22.1	188	6	BD006985 Artificialia
19	40	20.5	166	6	BD006983 Artificialia

20	12.8	1002	11	CNS06470	AI402010 T7 end of
21	11.8	182	6	BD006975	BD006975 Artificialia
22	10.8	127376	2	AC141672	AC141672 Apis mell
23	10.8	196217	2	AC131340	AC131340 Mus muscu
24	10.3	143255	2	AC015996	AC015996 Homo sapi
25	10.3	153887	9	AC093830	AC093830 Homo sapi
26	10.3	182737	10	AL772354	AL772354 Mouse DNA
27	10.3	199456	9	AC087664	AC087664 Homo sapi
28	10.3	235883	2	AC110956	AC110956 Rattus no
29	10.3	237468	2	AC108532	AC108532 Rattus no
30	10.3	309838	2	AC118433	AC118433 Rattus no
31	9.7	609	8	AY035415	AY035415 Typhlodin
32	9.7	1567	3	AF317926	AF317926 Typhlodin
33	9.7	1876	3	AF300529	AF300529 Diadabia
34	9.7	80449	2	AC114714	AC114714 Rattus no
35	9.7	110000	1	AE017308	AE017308 Continuation (5 of
36	9.7	129675	2	AC102950	AC102950 Homo sapi
37	9.7	145652	10	AL929262	AL929262 Mouse DNA
38	9.7	170606	10	AC102532	AC102532 Mus muscu
39	9.7	175436	2	AC122986	AC122986 Rattus no
40	9.7	179503	9	AL445468	AL445468 Human DNA
41	9.7	181527	10	AC140788	AC140788 Mus muscu
42	9.7	199814	3	CEY47D3A	AL117202 Caenorhab
43	9.7	204900	2	AC120282	AC120282 Rattus no
44	9.7	214074	2	AC148841	AC148841 Gorilla g
45	9.7	214455	2	AC118451	AC118451 Rattus no

ALIGNMENTS

RESULT 1
BD006986
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.

ACCESSION
BD006986
VERSION
JP 2001503249-A/55.
KEYWORDS
Saccharomyces cerevisiae (Baker's yeast)
SOURCE
Saccharomyces cerevisiae
ORGANISM
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.

REFERENCE
1 (bases 1 to 195)

AUTHORS
Hammer, K. and Janssen, P. R.

TITLE
Artificial promoter libraries for selected organisms and promoters
derived from such libraries

JOURNAL
Patent: JP 2001503249-A 55 13-MAR-2001;

COMMENT
PETER RUDAR JANSSEN
OS Saccharomyces cerevisiae (yeast)

PN JP 2001503249-A/55

PD 13-MAR-2001

PF 25-AUG-1997 JP 1998510287

PR 23-AUG-1996 DK 0886/96

PI KALIN HAMMER, PETER RUDAR JANSSEN

PC C12N15/09, C12N15/00

CC Strandness: Double;

CC Topology: Linear;

CC Key

FT promoter

location/Qualifiers

1..195

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 195; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. No. 6.7e-102;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGTACTAATGATAGTTATGCGCATGACGAGTCTTCTGACGTCTGAG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1607.43 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-55

Perfect score: 195

Sequence: 1 GAATTCGTCGACCAATGATGATCC 195

Scoring table: OLIGO_NUC

Word size: 0

Total number of hits satisfying chosen parameters: 68473088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	10.3	450	2	AM003938
C 2	20	10.3	775	9	CL582491
C 3	20	10.3	1145	6	CB565228
C 4	19	9.7	742	8	BH966805
C 5	19	9.7	815	8	CG350701
C 6	19	9.7	818	9	CG135560
C 7	19	9.7	909	9	CG084553
C 8	19	9.7	1270	3	CR678966
C 9	18	9.2	222	1	AA836473
C 10	18	9.2	228	6	CB449507
C 11	18	9.2	235	6	CB449507
C 12	18	9.2	235	6	CB449507
C 13	18	9.2	289	6	BY614706
C 14	18	9.2	407	8	CC153956
C 15	18	9.2	415	2	AM492605
C 16	18	9.2	437	8	BZ956398
C 17	18	9.2	438	8	AO382789
C 18	18	9.2	441	8	AZ702523
C 19	18	9.2	443	7	CR471577
C 20	18	9.2	451	9	CC094891
C 21	18	9.2	452	9	CC842207
C 22	18	9.2	469	8	BZ950076
C 23	18	9.2	502	8	BH004907
C 24	18	9.2	506	9	CR201997

C 25	18	9.2	518	8	BH402415
C 26	18	9.2	530	8	AZ284832
C 27	18	9.2	578	8	BP102832
C 28	18	9.2	579	8	AO710943
C 29	18	9.2	593	8	AZ255847
C 30	18	9.2	594	6	CA892299
C 31	18	9.2	621	8	BH022689
C 32	18	9.2	654	5	BM107202
C 33	18	9.2	659	9	BM161681
C 34	18	9.2	660	5	BM068182
C 35	18	9.2	664	9	BM121165
C 36	18	9.2	664	9	CR016365
C 37	18	9.2	675	2	BE744981
C 38	18	9.2	677	2	BE740303
C 39	18	9.2	681	9	CL750452
C 40	18	9.2	683	5	BM107288
C 41	18	9.2	691	9	BM107288
C 42	18	9.2	702	9	CR042981
C 43	18	9.2	716	9	BM184773
C 44	18	9.2	716	9	CC860315
C 45	18	9.2	723	9	BM187483

ALIGNMENTS

RESULT 1
LOCUS AM003938
DEFINITION w984f04.x1 NCI CGAP GC6 Homo sapiens CDNA clone IMAGE:2478751 3', mRNA sequence.
ACCESSION AM003938
VERSION AM003938.1 GI:5850854
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NCI-CCGAP
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at:
www.bio.lnll.gov/bbip/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC6"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; plasmid DNA from the normalized library NCI CGAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 172.243 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-56

Perfect score: 189
Sequence: 1 GAATTCGTGACTCACAAGAA.....TCGCTACCAATCATGATCC 189

Scoring table: OLIGO_NJC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2004as:*
12: Geneseq2004bs:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	189	2	AAV23228 Saccharom
2	76	40.2	195	2	AAV23222 Saccharom
3	68	36.0	167	2	AAV23218 Saccharom
4	68	36.0	177	2	AAV23215 Saccharom
5	68	36.0	195	2	AAV23225 Saccharom
6	68	36.0	195	2	AAV23227 Saccharom
7	68	36.0	195	2	AAV23220 Saccharom
8	68	36.0	195	2	AAV23176 Saccharom
9	65	34.4	193	2	AAV23223 Saccharom
10	58	30.7	179	2	AAV23221 Saccharom
11	51	27.0	195	2	AAV23230 Saccharom
12	51	27.0	195	2	AAV23229 Saccharom
13	48	25.4	191	2	AAV23219 Saccharom
14	46	24.3	191	2	AAV23217 Saccharom
15	40	21.2	166	2	AAV23224 Saccharom
16	34	18.0	188	2	AAV23226 Saccharom
17	23	12.2	182	2	AAV23216 Saccharom
18	18	9.5	786	13	ADR96610 M. tuberc
19	18	9.5	5198	6	AB132278 Human imm
20	18	9.5	6059	6	AB154343 Chemical1

21	18	9.5	16106	4	AAK67366 Human imm
22	18	9.5	16106	4	AA135957 Human mus
23	18	9.5	16106	8	ABX58945 CDNA enco
24	18	9.5	16106	12	ADJ29695 Human mus
25	18	9.5	110000	2	AAZ01425-08
26	18	9.5	110000	4	AA199682-02
27	18	9.5	110000	4	AA199683-02
28	17	9.0	43	12	ADH42842 Novel hum
29	17	9.0	498	12	ADL11516
30	17	9.0	576	6	ABQ44458
31	17	9.0	576	6	ABQ44459
32	17	9.0	1866	13	ADSA47979
33	17	9.0	1959	6	ADZ7607
34	17	9.0	2000	6	ADZ7607
35	17	9.0	2000	8	ADZ7607
36	17	9.0	2081	4	AB115437
37	17	9.0	2789	6	ABQ70842
38	17	9.0	3799	12	ADM97833
39	17	9.0	4159	4	AB110310
40	17	9.0	5326	4	AB110170
41	17	9.0	5572	6	ABJ32815
42	17	9.0	6000	6	ABJ33133
43	17	9.0	6160	6	ABK31273
44	17	9.0	6160	6	ABL70234
45	17	9.0	6163	6	ABN80118

ALIGNMENTS

RESULT 1	AAV23228 standard; DNA; 189 BP.
ID	AAV23228;
AC	28-JUN-1998 (first entry)
DT	Saccharomyces cerevisiae promoter Yp68.
DE	Saccharomyces cerevisiae; yeast promoter; optimise; spacer;
DM	artificial promoter library; gene expression; ds.
XX	Synthetic.
XX	Saccharomyces cerevisiae.
OS	Saccharomyces cerevisiae.
XX	Key
FT	Promoter
FT	Location/Qualifiers
FT	8.184
FT	/*tag= a
FT	/standard_name= "Yeast promoter"
XX	W09807846-Al.
XX	26-FEB-1998.
XX	25-AUG-1997; 97WO-DK000342.
XX	23-AUG-1996; 96DK-00000886.
XX	(JENS/) JENSEN P R.
XX	Jensen PR, Hammer K;
XX	WPI, 1998-179062/16.
XX	New artificial promoter libraries - containing consensus promoter
XX	sequences and variable spacers, used to generate promoters for optimising
XX	expression of genes.
XX	Claim 28; Page 68; 89pp; English.
XX	This is a Saccharomyces cerevisiae promoter sequence used in the
XX	construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 925.004 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-56

Perfect score: 189

Sequence: 1 GAATTCGACACTCAAGAGA.....TCGCTACCAATCATGATCC 189

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 2422767955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_bhg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189	100.0	189	6	BD006987 Artificialia
2	76	40.2	195	6	BD006981 Artificialia
3	68	36.0	167	6	BD006977 Artificialia
4	68	36.0	177	6	BD006974 Artificialia
5	68	36.0	195	6	BD006979 Artificialia
6	68	36.0	195	6	BD006984 Artificialia
7	68	36.0	195	6	BD006986 Artificialia
8	68	36.0	199	6	BD006934 Artificialia
9	65	34.4	193	6	BD006982 Artificialia
10	63	33.3	1844	8	SCYOL140W
11	63	33.3	9892	8	SCARGCDC
12	60	31.7	2479	8	SCYOL141W
13	58	30.7	179	6	BD006980 Artificialia
14	51	27.0	176	6	BD006989 Artificialia
15	51	27.0	195	6	BD006988 Artificialia
16	48	25.4	191	6	BD006978 Artificialia
17	46	24.3	191	6	BD006976 Artificialia
18	40	21.2	166	6	BD006983 Artificialia
19	34	18.0	188	6	BD006985 Artificialia

20	23	12.2	182	6	BD006975 Artificialia
c 21	21	11.1	127376	2	AC141672 Apis mell
c 22	21	11.1	159111	2	AC118864 Rattus no
c 23	21	11.1	196217	2	AC131340 Mus muscu
c 24	20	10.6	143255	2	AC015996 Homo sapi
c 25	20	10.6	153887	9	AC093830 Homo sapi
c 26	20	10.6	199456	9	AC087664 Homo sapi
c 27	20	10.6	237468	2	AC108532 Rattus no
c 28	20	10.6	309838	2	AC118433 Rattus no
c 29	19	10.1	1876	3	AF300529 Diadasta
c 30	19	10.1	110000	1	AB017308-4 Continuation (5 of
c 31	19	10.1	129675	2	AC102950 Homo sapi
c 32	19	10.1	177995	9	AC016700 Homo sapi
c 33	19	10.1	179485	9	AC005939 Homo sapi
c 34	19	10.1	179503	9	AL445468 Human DNA
c 35	19	10.1	181599	2	AC116896 Mus muscu
c 36	19	10.1	199814	3	CEY47D3A All11702 Caenorhab
c 37	19	10.1	204900	2	AC120282 Rattus no
c 38	19	10.1	214074	2	AC148841 Gorilla g
c 39	19	10.1	214455	2	AC118451 Rattus no
c 40	19	10.1	216113	2	AC116121 Mus muscu
c 41	18	9.5	584	5	GGCMP1 X12346 Chicken car
c 42	18	9.5	786	6	AR456027 Sequence
c 43	18	9.5	938	5	CR733067 Gallus ga
c 44	18	9.5	1362	5	CHKCMFX W97497 Gallus gall
c 45	18	9.5	2498	1	BTHKXA U03552 Bacillus th

ALIGNMENTS

RESULT 1
BD006987
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters
derived from such libraries.
ACCESSION
BD006987
VERSION
UP 2001503249-A/56
KEYWORDS
Saccharomyces cerevisiae (baker's yeast)
SOURCE
Saccharomyces cerevisiae
ORGANISM
Saccharomyces cerevisiae

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Saccharomyces cerevisiae (yeast)
PN UP 2001503249-A/56
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Saccharomyces cerevisiae (yeast)
PN UP 2001503249-A/56
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter

COMMENT
OS Saccharomyces cerevisiae (yeast)
PN UP 2001503249-A/56
PD 13-MAR-2001
PR 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC C12N15/09, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FT promoter

FEATURES
Source
Location/Qualifiers
1..189
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 189; DB 6; Length 189;
Best Local Similarity 100.0%; Pred. No. 2.4e-98;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAATTCGACACTCAAGAGTGGCGGGTGAAGCTGAGCCTGACACCTGGGCG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1557.97 seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-56

Perfect score: 189

Sequence: 1 GAATTCGACTCAGACAGAA.....TCGCTACCATCATGATCC 189

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6847908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database: EST:

1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_g881:
9: gb_g882:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	10.6	450	2	AM003938
C 2	20	10.6	1145	6	CB565228
C 3	19	10.1	328	4	BC302359
C 4	19	10.1	491	8	BZ881549
C 5	19	10.1	522	1	AA062245
C 6	19	10.1	594	4	BG181949
C 7	19	10.1	640	9	CL337192
C 8	19	10.1	686	4	BG354907
C 9	19	10.1	742	8	BH966805
C 10	19	10.1	803	2	BF789111
C 11	19	10.1	857	5	BZ881551
C 12	18	9.5	212	5	BM511685
C 13	18	9.5	228	6	CB449507
C 14	18	9.5	235	6	CB440185
C 15	18	9.5	235	6	CB440185
C 16	18	9.5	303	1	AA874226
C 17	18	9.5	375	5	AA874226
C 18	18	9.5	407	8	BH692884
C 19	18	9.5	407	8	CC153956
C 20	18	9.5	451	8	CC094891
C 21	18	9.5	452	9	CC842207
C 22	18	9.5	502	8	AZ060925
C 23	18	9.5	502	8	BH004907
C 24	18	9.5	518	5	BM525876

25	18	9.5	536	8	BH027087
26	18	9.5	548	5	BU386212
27	18	9.5	578	5	BP102832
28	18	9.5	593	8	A2255847
29	18	9.5	621	1	AJ658080
30	18	9.5	648	8	A2417864
31	18	9.5	653	2	BB640068
32	18	9.5	654	2	BM107202
33	18	9.5	660	5	BM068182
34	18	9.5	676	6	CD438002
35	18	9.5	679	6	BU257424
36	18	9.5	679	6	BY751035
37	18	9.5	681	9	CT750452
38	18	9.5	683	5	BM107288
39	18	9.5	704	8	BH560191
40	18	9.5	716	4	BI756006
41	18	9.5	716	9	CC860315
42	18	9.5	717	9	CR042952
43	18	9.5	727	9	AG611513
44	18	9.5	760	9	CC874751
45	18	9.5	775	8	BH439976

ALIGNMENTS

RESULT 1
AM003938/c
LOCUS
DEFINITION
w884f04.x1 NCI_CGAP_G66 Homo sapiens cDNA clone IMAGE:2478751 3',
ACCESSION
AM003938
VERSION
AM003938.1 GI:5850854
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 450)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LML at:
www.bio.lml.gov/ldbr/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/ld_host="DH10B"
/clone_lib="NCI CGAP G66"
/note="Vector: pTR73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I. Site 2: Bco RI. Plasmid DNA from the normalized library NCI CGAP G64 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LML at:
www.bio.lml.gov/ldbr/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/ld_host="DH10B"
/clone_lib="NCI CGAP G66"
/note="Vector: pTR73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I. Site 2: Bco RI. Plasmid DNA from the normalized library NCI CGAP G64 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES

source
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LML at:
www.bio.lml.gov/ldbr/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 450
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2478751"
/tissue_type="pooled germ cell tumors"
/ld_host="DH10B"
/clone_lib="NCI CGAP G66"
/note="Vector: pTR73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I. Site 2: Bco RI. Plasmid DNA from the normalized library NCI CGAP G64 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 177.711 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-57

Perfect score: 195
Sequence: 1 GAATCTGCTGACTGCTGGA.....TCGCTACCATCATGATCC 195

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	195	2	AAV23229 Saccharom
2	53	27.2	176	2	AAV23230 Saccharom
3	51	26.2	167	2	AAV23218 Saccharom
4	51	26.2	177	2	AAV23215 Saccharom
5	51	26.2	179	2	AAV23221 Saccharom
6	51	26.2	189	2	AAV23228 Saccharom
7	51	26.2	193	2	AAV23223 Saccharom
8	51	26.2	195	2	AAV23225 Saccharom
9	51	26.2	195	2	AAV23222 Saccharom
10	51	26.2	195	2	AAV23227 Saccharom
11	51	26.2	195	2	AAV23220 Saccharom
12	51	26.2	199	2	AAV23176 Saccharom
13	48	24.6	191	2	AAV23219 Saccharom
14	46	23.6	191	2	AAV23217 Saccharom
15	40	20.5	166	2	AAV23224 Saccharom
16	23	11.8	182	2	AAV23216 Saccharom
17	21	10.8	188	2	AAV23226 Saccharom
18	19	9.7	566	6	ABQ53718 Oligonuc1
19	19	9.7	566	6	ABQ53719 Oligonuc1
20	16	9.2	648	6	ABQ16662 Oligonuc1

c	21	18	9.2	648	6	ABQ16663	Abq16663 Oligonuc1
c	22	18	9.2	737	6	ABQ38638	Abq38638 Oligonuc1
c	23	18	9.2	737	6	ABQ38639	Abq38639 Oligonuc1
c	24	18	9.2	5198	6	ABJ32278	Abj32278 Human imm
c	25	18	9.2	6059	6	ABL54343	AbL54343 Chemical
c	26	18	9.2	6104	4	AA546295	AA546295 Tumour su
c	27	18	9.2	6104	6	ABJ32296	Abj32296 Human imm
c	28	18	9.2	6104	10	ADB54103	AdB54103 Preterate
c	29	18	9.2	6104	13	AD889259	Ad889259 Oligonuc1
c	30	18	9.2	11812	4	AA545502	AA545502 Chemical
c	31	18	9.2	11812	4	AA546742	AA546742 Tumour su
c	32	18	9.2	11812	6	ABL34119	AbL34119 Human imm
c	33	18	9.2	11812	6	ABK28432	AbK28432 DNA trans
c	34	18	9.2	96596	9	ADA02864	Ada02864 Human ITP
c	35	18	9.2	96596	10	ADB72802	AdB72802 Human ITP
c	36	18	9.2	96596	10	ADC85343	AdC85343 Mouse ITP
c	37	18	9.2	96596	12	ADM74459	AdM74459 Human car
c	38	18	9.2	110000	2	AA201425	AA201425 (9 of
c	39	17	8.7	43	12	ADH42842	AdH42842 Novel hum
c	40	17	8.7	212	3	AA133088	AA133088 Human sec
c	41	17	8.7	216	6	ABK76865	AbK76865 Bacillus
c	42	17	8.7	332	4	AA186773	AA186773 Human pol
c	43	17	8.7	418	9	ACH37565	ACH37565 Human end
c	44	17	8.7	436	3	AAC09917	AAc09917 Human sec
c	45	17	8.7	443	9	ACH50822	ACH50822 Human mam

ALIGNMENTS

RESULT 1	AAV23229
ID	AAV23229 standard; DNA; 195 BP.
XX	AAV23229;
XX	28-UTL-1998 (first entry)
XX	Saccharomyces cerevisiae promoter Yp89.
XX	Saccharomyces cerevisiae; Yeast promoter; optimise; spacer;
KW	artificial promoter library; gene expression; db.
XX	Synthetic.
OS	Saccharomyces cerevisiae.
OS	Saccharomyces cerevisiae.
FT	Key
FT	Location/Qualifiers
FT	8..190
FT	/*tag= a
FT	/standard_name= "Yeast promoter"
XX	W09807846-A1.
XX	26-FEB-1998.
XX	25-AUG-1997; 97WO-DK000342.
XX	23-AUG-1996; 96DK-00000886.
XX	(JENS/) JENSEN P R.
XX	Jensen PR, Hammer K;
XX	WPI; 1998-179062/16.
PT	New artificial promoter libraries - containing consensus promoter
PT	sequences and variable spacers, used to generate promoters for optimising
PT	expression of genes.
XX	Claim 28; Page 68-69; 89pp; English.
CC	This is a Saccharomyces cerevisiae promoter sequence used in the
CC	construction of an artificial promoter library of the invention. The

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 954.369 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-57

Perfect score: 195
Sequence: 1 GAATGCTGACTGACTGCGA.....TCGCTACCATCATGATCC 195

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : GenEmbl.*

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_sts:*
12: gb_by:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	195	6	BD006988 Artificia
2	53	27.2	176	6	BD006989 Artificia
3	51	26.2	167	6	BD006977 Artificia
4	51	26.2	177	6	BD006974 Artificia
5	51	26.2	179	6	BD006980 Artificia
6	51	26.2	189	6	BD006987 Artificia
7	51	26.2	193	6	BD006982 Artificia
8	51	26.2	195	6	BD006979 Artificia
9	51	26.2	195	6	BD006981 Artificia
10	51	26.2	195	6	BD006986 Artificia
11	51	26.2	195	6	BD006984 Artificia
12	51	26.2	199	6	BD006934 Artificia
13	48	24.6	191	6	BD006978 Artificia
14	46	23.6	191	6	BD006976 Artificia
15	46	23.6	1844	8	BD006940W
16	46	23.6	9892	8	SCARGCDC
17	43	22.1	2479	8	SCYOL141W
18	40	20.5	166	6	BD006983 Artificia
19	33	16.9	188	6	BD006985 Artificia

20	23	11.8	182	6	BD006975	BD006975 Artificia
21	23	11.8	1002	11	CNS06JTO	AL402010 T7 end of
22	21	10.8	1876	3	AF300529	AF300529 Diadadia
23	21	10.8	11000	1	BR008798_11	Continuation (12 o
24	21	10.8	127376	2	AC141672	AC141672 Apis mell
25	21	10.8	196217	2	AC131340	AC131340 Mus muscu
26	20	10.3	1503	8	CURPL25	X05919 Candida uci
27	20	10.3	131189	2	AC101756	AC101756 Mus muscu
28	20	10.3	143255	2	AC091596	AC091596 Homo sapi
29	20	10.3	153887	9	AC093830	AC093830 Homo sapi
30	20	10.3	199456	9	AC087664	AC087664 Homo sapi
31	20	10.3	206478	10	AL928693	AL928693 Mouse DNA
32	20	10.3	237468	2	AC108532	AC108532 Rattus no
33	20	10.3	309893	3	AC118433	AC118433 Rattus no
34	19	9.7	333	3	AY228755	AY228755 Parametru
35	19	9.7	110000	2	AF546187_2	Continuation (3 of
36	19	9.7	110000	2	AF546187_2	Continuation (3 of
37	19	9.7	121624	2	AC148843	AC148843 Oryzomys
38	19	9.7	174706	2	AC150618	AC150618 Callithrix
39	19	9.7	179590	10	AC113293	AC113293 Mus muscu
40	19	9.7	197219	2	AC113278	AC113278 Mus muscu
41	19	9.7	199814	3	CEV47D3A	AL117202 Ctenorhab
42	19	9.7	204900	2	AC120282	AC120282 Rattus no
43	19	9.7	210156	2	AC136829	AC136829 Rattus no
44	19	9.7	214074	2	AC148841	AC148841 Gorilla g
45	19	9.7	214455	2	AC118451	AC118451 Rattus no

ALIGNMENTS

RESULT 1
LOCUS BD006988 195 bp DNA linear PAT 31-JUN-2002
DEFINITION Artificial promoter libraries for selected organisms and promoters
derived from such libraries.

ACCESSION BD006988.1 GI:18635359
VERSION JP 2001503249-A/57
KEYWORDS Saccharomyces cerevisiae (baker's yeast)
SOURCE Saccharomyces cerevisiae
ORGANISM Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE 1 (bases 1 to 195)
AUTHORS Hammer, K. and Janssen, P. R.
TITLE Artificial promoter libraries for selected organisms and promoters
derived from such libraries
JOURN. Patent: JP 2001503249-A 57 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT OS Saccharomyces cerevisiae (yeast)
PN JP 2001503249-A/57
PD 13-MAR-2001
PF 25-AUG-1997 JP 1998510287
PR 23-AUG-1996 DK 0886/96
PI KALIN HAMMER, PETER RUDAR JANSSEN
PC Cl2N15/09, Cl2N15/00
CC Strandedness: Double;
CC Topology: linear;
FH Key Location/Qualifiers
FT promoter (8)..(190).
Location/Qualifiers

FEATURES
source 1..195
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN
Query Match 100.0%; Score 195; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.7e-102; Indels 0; Gaps 0;
Matches 195; Conservative 0

1 GAATGCTGACTGACTGCGAGATTGGTTTACGATTAGATGCGCGGACACCGGGG 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1607.43 Seconds

(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657B-57

Sequence: 1 GAATTCGTGACTACTCGGA.....TCGCTACCAATGATGATCC 195

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	10.3	775	9	CL582491 OB_Ba004
C 2	20	10.3	1145	6	CB565228 AGENCOURT
C 3	19	9.7	104	5	BX700116 BX700116
C 4	19	9.7	471	7	CO376822 ta143b11.
C 5	19	9.7	526	1	AT244109 qj63e07.x
C 6	19	9.7	600	9	CC640310 OCGY72TH
C 7	19	9.7	853	9	CG928655 MBELN29TR
C 8	19	9.7	845	9	CG205931 PUTBG40TB
C 9	19	9.7	935	5	CG937986 MBELV06TR
C 10	19	9.7	935	5	BQ930457 AGENCOURT
C 11	18	9.2	156	7	CO340510 EP10805.3
C 12	18	9.2	228	6	CB449507 703710.MA
C 13	18	9.2	232	1	AV312820 AV312820
C 14	18	9.2	235	6	CB440185 690440.MA
C 15	18	9.2	235	6	CB451827 706622.MA
C 16	18	9.2	267	7	CN777610 ta65h01.
C 17	18	9.2	296	5	BW505456 BM505456
C 18	18	9.2	296	7	CO219595 WS0107.B2
C 19	18	9.2	351	5	BP519849 BP519849
C 20	18	9.2	363	5	BP023739 BP023739
C 21	18	9.2	407	8	CC153956 CSU-K34.1
C 22	18	9.2	425	7	CN623429 ta606h08
C 23	18	9.2	443	7	CR471577 CR471577
C 24	18	9.2	451	8	CC094891 CSU-K34.1

C 25	18	9.2	452	9	CC842207 ND1.97010
C 26	18	9.2	467	7	CO223605 WS01019.B
C 27	18	9.2	483	7	CO276705 EK121859.
C 28	18	9.2	485	6	CB955777 AGENCOURT
C 29	18	9.2	502	8	BH004907 BM34106
C 30	18	9.2	546	5	BM34106 BM34106
C 31	18	9.2	552	6	CB247053 UI-M-F10-
C 32	18	9.2	563	5	BM332299 BM332299
C 33	18	9.2	576	4	BM654928 BM654928
C 34	18	9.2	575	4	BM017498 BM017498
C 35	18	9.2	577	7	CN556503 ta637d08.
C 36	18	9.2	578	5	BP102832 BP102832
C 37	18	9.2	590	9	CE257358 tigr-gss-
C 38	18	9.2	593	8	A2255847 RPTC1-23-1
C 39	18	9.2	655	5	BM363217 BM363217
C 40	18	9.2	664	9	BM121165 BM121165
C 41	18	9.2	667	7	CN768148 CN768148
C 42	18	9.2	669	5	BM287205 BM287205
C 43	18	9.2	681	9	CL750452 OR_BA011
C 44	18	9.2	688	4	BM579586 BM579586
C 45	18	9.2	688	5	BM035245 BM035245

ALIGNMENTS

RESULT 1
LOCUS CL582491/c
DEFINITION OB_Ba0047L19.1 OB_Ba Oryza brachyantha genomic clone
VERSION OB_Ba0047L19.1
KEYWORDS CL582491
SOURCE CL582491.1 GI:48540712
ORGANISM GSS.
Oryza brachyantha
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 775)
Kim,H., Yu,X., Stum,D., Yost,D., Rao,K., Luo,M., Jetty,R.,
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.
OMAP Project
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TTA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Insert Length: 14500 Std Error: 0.00
Plate: 0047 row: 1 column: 19
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES
Location/Qualifiers
1..775
/organism="Oryza brachyantha"
/mol_type="genomic DNA"
/db_xref="taxon:4533"
/clone="OB_Ba0047L19"
/issue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OB_Ba"
/note="Vector: pAG1BAC1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 10.3%; Score 20; DB 9; Length 775;
Best Local Similarity 100.0%; Pred. No. 3.1;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 09:46:05 ; Search time 160.396 Seconds
(without alignments)
6495.658 Million cell updates/sec

Title: US-09-242-657B-58

Sequence: 1 GAATCGTACATCATCTAGT.....TCGCTACCAATCATGATCC 176

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	100.0	176	2	AAV23230 Saccharom
2	53	30.1	195	2	AAV232329 Saccharom
3	51	29.0	167	2	AAV23218 Saccharom
4	51	29.0	177	2	AAV23215 Saccharom
5	51	29.0	179	2	AAV23221 Saccharom
6	51	29.0	189	2	AAV23228 Saccharom
7	51	29.0	193	2	AAV23223 Saccharom
8	51	29.0	195	2	AAV23225 Saccharom
9	51	29.0	195	2	AAV23222 Saccharom
10	51	29.0	195	2	AAV23227 Saccharom
11	51	29.0	195	2	AAV23220 Saccharom
12	51	29.0	199	2	AAV23176 Saccharom
13	48	27.3	191	2	AAV23219 Saccharom
14	46	26.1	191	2	AAV23217 Saccharom
15	40	22.7	166	2	AAV23224 Saccharom
16	23	13.1	182	2	AAV23216 Saccharom
17	18	10.2	171	12	ADG64222 Y chromos
18	18	10.2	498	6	ACH50638 Human mam
19	18	10.2	566	6	ABG53718 Oligonuc1
20	18	10.2	566	6	ABG53719 Oligonuc1

21	18	10.2	645	13	ADG53986 Novel can
22	18	10.2	712	8	ABZ51234 Aspergill
23	18	10.2	1253	2	AAV34241 Human sec
24	18	10.2	1253	8	ACD08112 CDNA enco
25	18	10.2	1428	3	AACT78073 Human can
26	18	10.2	2804	10	ADD48414 Rat gene
27	18	10.2	3234	8	ABZ76275 Notch-int
28	18	10.2	3308	12	ADG64477 Novel hum
29	18	10.2	5198	6	ABL32278 Human imm
30	18	10.2	6059	6	ABL54343 Chemical1
31	18	10.2	8070	13	ACN40496 Tumour-as
32	18	10.2	8091	2	AAV57001 Human Not
33	18	10.2	8091	6	ABZ34974 Human gen
34	18	10.2	8091	10	ADL13816 Osteoarth
35	18	10.2	8091	12	ADH34620 Notch hom
36	18	10.2	8091	12	ADJ75112 Marker ge
37	18	10.2	8091	13	ADN48549 Human Not
38	18	10.2	8091	13	ADP23495 PRO polyp
39	18	10.2	8091	13	ADR83501 Human Not
40	18	10.2	41150	10	ADL13819 Osteoarth
41	18	10.2	44348	12	ADN48556 Human Not
42	18	10.2	95109	6	ABQ99654 Human mem
43	18	10.2	110000	2	AAZ01425 08 Continuation (9 of
44	17	9.7	43	12	ADH42842 Novel hum
45	17	9.7	188	2	AAV23226 Saccharom

ALIGNMENTS

RESULT 1	AAV23230	AAV23230 standard; DNA, 176 BP.
XX	AAV23230;	
AC	28-JUL-1998	(first entry)
XX		
DE	Saccharomyces cerevisiae promoter Yp96.	
XX	Saccharomyces cerevisiae; yeast promoter; optimise; spacer;	
KW	artificial promoter library; gene expression; ds.	
XX		
OS	Synthetic.	
XX	Saccharomyces cerevisiae.	
XX		
Key	Location/Qualifiers	
FT	promoter	8..171
FT		/*tag= a
FT		/standard_name= "Yeast promoter"
XX		
PD	WO9807846-A1.	
XX	26-FEB-1998.	
XX		
PF	25-AUG-1997;	97WO-DK000342.
XX		
PR	23-AUG-1996;	96DK-00000886.
XX		
PA	(JENSEN) JENSEN P R.	
XX		
PI	Jensen PR, Hammer K;	
XX		
DR	WPI, 1998-179062/16.	
XX		
PT	New artificial promoter libraries - containing consensus promoter	
PT	sequences and variable spacers, used to generate promoters for optimising	
XX	expression of genes.	
XX		
PS	Claim 28; Page 69; 89pp; English.	
CC	This is a Saccharomyces cerevisiae promoter sequence used in the	
CC	construction of an artificial promoter library of the invention. The	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 12:37:20 ; Search time 861.379 Seconds
(without alignments)
9900.540 Million cell updates/sec

Title: US-09-242-657B-58

Perfect score: 176
Sequence: 1 GAATCGTACTCATCTAGT.....TCGCTACCATCATGATCC 176

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sv: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	100.0	176	6	BD006989
2	53	30.1	195	6	BD006988
3	51	29.0	167	6	BD006977
4	51	29.0	177	6	BD006974
5	51	29.0	179	6	BD006980
6	51	29.0	189	6	BD006987
7	51	29.0	193	6	BD006982
8	51	29.0	195	6	BD006979
9	51	29.0	195	6	BD006981
10	51	29.0	195	6	BD006984
11	51	29.0	195	6	BD006986
12	51	29.0	199	6	BD006934
13	48	27.3	191	6	BD006978
14	46	26.1	191	6	BD006976
15	46	26.1	1844	8	SCYOL140M
16	46	26.1	9892	8	SCARGCDD
17	43	24.4	2479	8	SCYOL141W
18	40	22.7	166	6	BD006983
19	33	18.8	188	6	BD006985

20	23	13.1	182	6	BD006975	BD006975
21	22	12.5	1876	3	AF300529	AF300529
22	21	11.9	127376	2	AC141672	AC141672
23	21	11.9	196217	2	AC131340	AC131340
24	21	11.9	218905	10	AC087772	AC087772
25	21	11.9	220811	10	AC087795	AC087795
26	21	11.9	225541	2	AC126700	AC126700
27	21	11.9	261093	10	AC027740	AC027740
28	20	11.4	484	6	AR504595	AR504595
29	20	11.4	484	6	AR519877	AR519877
30	20	11.4	143255	6	AC015996	AC015996
31	20	11.4	153887	9	AC093830	AC093830
32	20	11.4	199456	9	AC087664	AC087664
33	20	11.4	237468	2	AC108532	AC108532
34	20	11.4	309838	2	AC118433	AC118433
35	19	10.8	1099	3	PF025814	PF025814
36	19	10.8	110000	1	AE017308-4	AE017308-4
37	19	10.8	114848	9	AC018664	AC018664
38	19	10.8	115723	9	AL670662	AL670662
39	19	10.8	119916	9	AF241734	AF241734
40	19	10.8	159012	2	AC119414	AC119414
41	19	10.8	163215	9	AC108881	AC108881
42	19	10.8	163443	2	AC006280	AC006280
43	19	10.8	169002	2	AC102427	AC102427
44	19	10.8	187893	9	AC145905	AC145905
45	19	10.8	196149	2	AC004709	AC004709

ALIGNMENTS

RESULT 1
BD006989
LOCUS
DEFINITION
Artificial promoter libraries for selected organisms and promoters derived from such libraries.

ACCESSION
BD006989
VERSION
JP 2001503249-A/58
KEYWORDS
Saccharomyces cerevisiae (Baker's yeast)
SOURCE
Saccharomyces cerevisiae
ORGANISM
Saccharomyces cerevisiae

REFERENCE
1 (bases 1 to 176)
AUTHORS
Hammer, K. and Janssen, P.R.
TITLE
Artificial promoter libraries for selected organisms and promoters derived from such libraries
JOURNAL
Patent: JP 2001503249-A 58 13-MAR-2001;
PETER RUDAR JANSSEN

COMMENT
OS
Saccharomyces cerevisiae (yeast)
PN
JP 2001503249-A/58
PD
13-MAR-2001
PF
25-AUG-1997 JP 1998510287
PR
23-AUG-1996 DK 0886/96

PI
KALIN HAMMER, PETER RUDAR JANSSEN
PC
C12N15/09, C12N15/00
CC
Strandness: Double;
Topology: Linear;
FH
Key
FT
promoter
location/Qualifiers
(8). (111).

FEATURES
source
1..176
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"

ORIGIN

Query Match 100.0%; Score 176; DB 6; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.3e-90;
Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GAATCGTACTCATCTAGTACAGAGAGCCGTATCTGTGTCACCAAGGAGAAA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2005, 20:41:42 ; Search time 1450.81 Seconds
(without alignments)
4617.633 Million cell updates/sec

Title: US-09-242-657b-58

Perfect score: 176
Sequence: 1 GAATTCGACTCATCTACT.....TCGCTACCAATCATGATCC 176

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_g861:
9: gb_g862:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	11.4	156	7	CO340510 EP10805.3
2	20	11.4	483	7	CO216705 EX121859
3	20	11.4	547	6	CF035245 OCG15h02
4	20	11.4	1145	6	CB565228 AGENCOURT
5	19	10.8	158	5	BU496440 PESTROADS
6	19	10.8	195	5	BO597017 PESTROADS
7	19	10.8	291	7	CV231616 MS0195.B2
8	19	10.8	369	7	CF404266 CSECS024B
9	19	10.8	429	5	BQ792641 EST 8361
10	19	10.8	476	6	CA815842 CA12E1301
11	19	10.8	482	6	CA817911 CA12E1301
12	19	10.8	516	6	CA923437 MTU7CL.P1
13	19	10.8	538	7	CV252179 MS0119.B2
14	19	10.8	547	7	CV084633 FAMU_USDA
15	19	10.8	556	6	CD798074 EST 14455
16	19	10.8	589	8	BH697204 BOMAV80TR
17	19	10.8	612	8	BH516587 BOGHL30TR
18	19	10.8	617	6	CD004321 VVA011A08
19	19	10.8	617	7	CF210418 CAB20006
20	19	10.8	647	7	CF515034 CAP0001.I
21	19	10.8	650	5	BO796247 EST 5185
22	19	10.8	670	7	CV088007 FAMU_USDA
23	19	10.8	677	7	CV245999 WS0259.B2
24	19	10.8	694	9	CL911429 OA_Aba001

25	19	10.8	708	8	B2472842 BONHM30TR
26	19	10.8	716	9	BX173581 Danilo rer
27	19	10.8	718	6	CA288726 SCEOSD207
28	19	10.8	737	7	CF202060 RR890915N
29	19	10.8	746	7	CF514855 CAP0001.I
30	19	10.8	752	6	CD003727 VVA011A08
31	19	10.8	753	7	CV098283 FAMU_USDA
32	19	10.8	762	7	CF210502 CAB20006
33	19	10.8	766	7	CV234551 MS01215.B
34	19	10.8	778	6	CA816005 CA12E1301
35	19	10.8	778	7	CF211942 CGF100066
36	19	10.8	780	8	BH660637 BOHVP35TR
37	19	10.8	790	1	AJ812966 AJ812966
38	19	10.8	799	7	CNS20439 G00107.B3
39	19	10.8	805	9	CG414379 ZM8Bb028
40	19	10.8	808	7	CF211795 CGF100066
41	19	10.8	852	7	CF211885 CGF100066
42	19	10.8	864	7	CV278137 WS0145.B2
43	19	10.8	876	7	CK097374 UB49DPA01
44	19	10.8	918	7	CV260369 MS02013.B
45	19	10.8	943	9	CG905601 ZM8Bb051

ALIGNMENTS:

RESULT 1
CO340510
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Plate: EP.108 Row: A Column: 5
High quality sequence stop: 154.
Location/Qualifiers
1. 156
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="EP10805"
/note="Organ: mixed stage embryos, imaginal disks, and adult heads; Vector: pCDNA-SK+; Site 1: NotI; Site 2: XhoI; Random primed, normalized library from mixed stage embryos, imaginal disks, and adult heads. Subset of Exelixis FlyTag CK01 clones sequenced from 3' end"

Query Match 11.4%; Score 20; DB 7; Length 156;
Best Local Similarity 100.0%; Pred. No. 4.4; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
122 TTAATTTTTCGTTGTT 141
88 TTAATTTTTCGTTGTT 107